R. Semper
Johns Hopkins U.

QUANTUM MECHANICS

QUANTUM MECHANICS

F. MANDL, M.A., D.Phil.
Department of Theoretical Physics,
University of Manchester

SECOND EDITION

LONDON
BUTTERWORTHS
1957

ENGLAND: BUTTERWORTH & CO. (PUBLISHERS) LTD.
LONDON: 88 Kingsway, W.C.2
AUSTRALIA: BUTTERWORTH & CO. (AUSTRALIA) LTD.
SYDNEY: 20 Loftus Street
MELBOURNE: 473 Bourke Street
BRISBANE: 240 Queen Street
CANADA: BUTTERWORTH & CO. (CANADA) LTD.
TORONTO: 1367 Danforth Avenue, 6
NEW ZEALAND: BUTTERWORTH & CO. (NEW ZEALAND) LTD.
WELLINGTON: 49/51 Ballance Street
AUCKLAND: 35 High Street
SOUTH AFRICA: BUTTERWORTH & CO. (SOUTH AFRICA) LTD.
DURBAN: 33/35 Beach Grove
U.S.A.: BUTTERWORTH INC.
WASHINGTON, D.C.: 7300 Pearl Street, 200 14

1st Edition, 1954
2nd Edition, 1957
2nd Impression, 1958
3rd Impression, 1960
4th Impression, 1961
5th Impression, 1962
6th Impression, 1964
7th Impression, 1966

©
Butterworth & Co. (Publishers) Ltd.
1957

Set in Monotype Baskerville type
Made and printed in Great Britain by William Clowes and Sons, Limited
London and Beccles

CONTENTS

PAGE

PREFACE TO THE SECOND EDITION viii

PREFACE TO THE FIRST EDITION ix

I. MATHEMATICAL INTRODUCTION 1

1. Vectors 1
2. Vibrating string 5
3. Mathematical generalizations 8
4. Continuous spectra 14

Exercises I 20

II. WAVE-MECHANICAL CONCEPTS 22

5. The Schrödinger wave equation: derivation 23
6. The Schrödinger wave equation: discussion 25
7. Energy eigenfunctions: theory 31

Exercises II 34

III. ENERGY EIGENFUNCTIONS: EXAMPLES 36

8. One-dimensional potential barrier 36
9. One-dimensional square well potential with rigid walls 38
10. Free particle wave functions and three-dimensional square well potential 40
11. Linear harmonic oscillator 43
12. Spherically symmetric potentials in three dimensions .. 47
13. Spherical harmonics 49
14. Angular momentum 54

Exercises III 56

IV. GENERAL PRINCIPLES OF QUANTUM MECHANICS 58

15. Résumé of classical mechanics 58
16. Interpretative postulates of quantum mechanics .. 60
17. Operators with continuous spectra 70
18. Functions of operators 71
19. Simultaneous measurability of observables 74
20. Uncertainty principle 82

Exercises IV 88

PAGE
V. MATRIX MECHANICS 90
21. Matrix representation of wave functions and operators .. 90
22. Properties of matrix operators 93
(*a*) Matrix multiplication 93
(*b*) Adjoint matrices 95
(*c*) Change of basis 95
(*d*) Eigenvalue problems 97
23. The equations of motion of operators 100
24. The Dirac notation for wave functions and operators .. 101
Exercises V 102

VI. SYSTEMS OF MANY PARTICLES 104
25. Angular momentum 104
26. Spin 110
27. Addition of angular momenta 116
28. Systems of identical particles: Pauli principle 125
Exercises VI 130

VII. TIME-INDEPENDENT PERTURBATION THEORY 131
29. First-order perturbation: non-degenerate case 131
30. First-order perturbation: degenerate case 134
31. Simple examples on perturbation theory 136
Example 1: Ground state of the helium atom .. 136
Example 2: Two identical particles 137
Example 3: Normal Zeeman effect 140
Exercises VII 141

VIII. COLLISION PROCESSES 142
32. Elastic scattering by a fixed centre of force 142
33. The integral equation of the scattering problem .. 145
34. Born approximation 149
Example 1: Scattering by a square well potential .. 151
Example 2: Scattering by a screened Coulomb potential 151
35. Laboratory and centre of mass coordinate systems .. 153
36. Inelastic scattering processes 158
37. Time-dependent approach to collision problems .. 161
38. Partial waves 166
Exercises VIII 175

PAGE
IX. INTRODUCTION TO GROUP-THEORETICAL IDEAS 177
39. Groups of transformations 177
40. Rotation operators 180
41. Invariants of groups of transformations 181
42. Spin 183
43. The representation of groups 186
44. Perturbations 191
45. Selection rules 195
Exercises IX 201

X. RELATIVISTIC THEORY OF THE ELECTRON 203
46. The Dirac equation for a free particle 203
47. The spin and magnetic moment of the electron .. 209
48. Plane wave solutions of the Dirac equation 212
49. Relativistic invariance 216
Exercises X 222

HINTS FOR SOLVING THE EXERCISES 224

APPENDIX A: VECTOR ANALYSIS 261

BIBLIOGRAPHY 263

INDEX 265

PREFACE TO THE SECOND EDITION

In this second edition the approach and scope of the first edition have not been altered. In particular, no attempt has been made to deal more comprehensively with standard applications of quantum mechanics which are treated well in existing textbooks. However, following several suggestions, two new topics are included.

I have added a section on partial wave analysis in scattering theory. This subject is so important in atomic, nuclear and high-energy physics that it seemed desirable to include it. A new chapter at the end of the book is devoted to the Dirac equation. This contains a discussion of the relativistic invariance of the theory which underlies the construction of interactions in beta-decay and field theories generally. In this chapter some knowledge of the special theory of relativity is presupposed on the part of the reader. In addition, a number of smaller changes and corrections, for some of which I am indebted to readers, have been incorporated. Of these I should mention the redefinition of spherical harmonics (section 13) so that their phases now agree with those used almost exclusively in angular momentum problems.

I should again like to express my gratitude to my friend Dr. Handel Davies, of Christ Church, with whom I discussed the proposed changes and who read and criticized the new material; also to my wife, Betty, for assisting me throughout the preparation of this edition.

F.M.

Harwell, Berks
April, 1957

PREFACE TO THE FIRST EDITION

In the autumn of 1952 I gave a course of lectures on quantum mechanics to experimentalists at A.E.R.E., Harwell, in which I developed the theory in such a way as to bring out the unifying mathematical scheme underlying quantum mechanics. This book, based on these lectures, uses the same approach.

In the first five chapters the mathematical formalism of quantum mechanics is systematically developed, with particular attention to its physical interpretation. Mathematical rigour or generality are not aimed at, the stress being on a strongly geometrical formulation of the mathematics, given in considerable detail and kept as simple as possible. The mathematical methods underlying quantum mechanics are developed in Chapter I, new physical concepts being introduced only in Chapter II. The simple but important applications of Chapter III should make the reader familiar with these concepts which, in Chapter IV, are extended to the general scheme of quantum mechanics. Questions of the measurability of one or more observables are here discussed in much greater detail than is customary in all but very advanced textbooks, but I believe that an understanding of these problems is essential to an understanding of quantum mechanics. Matrix mechanics is treated in Chapter V, not using correspondence arguments (or mere formal analogies) from classical mechanics, which are difficult to appreciate, but as a particular representation of the general theory, as it arises in practice.

The use of the general theory in solving specific problems is illustrated in the remaining four chapters. Since many existing books deal satisfactorily with such applications, I have restricted myself to a few topics, of particular importance in many branches of applied quantum mechanics. Chapter VI deals with angular momentum, spin and symmetry properties of systems of particles. These questions are so important in atomic and nuclear physics that I have treated them in considerable detail. Perturbation theory, one of the most useful methods for handling complicated problems, is treated in Chapter VII, while Chapter VIII is devoted to scattering problems. The latter form such a vast subject that I have only been able to deal with some principal aspects; some topics, such as the phase-shift analysis of scattering problems, have been omitted, as they are treated well in many existing books. In Chapter IX the basic ideas of group theory are developed and applied to perturbation theory, and the derivation of conservation laws and selection rules. The introduction of group-theoretical concepts seemed desirable since they lead to a uniform treatment of many diverse problems, and occur frequently in the literature.

At the end of each chapter, examples are given, which form an integral part of the book as reference is made to them and a knowledge of them is assumed later. A section giving detailed hints for their solutions is therefore included, and the reader is advised to study this, even if he does not work through the examples himself.

The knowledge presupposed on the part of the reader is kept to a minimum. Apart from some classical physics, calculus and vector analysis, he is assumed to be familiar with the qualitative concepts of quantum theory, such as the wave nature of matter, and their experimental basis. It is thus hoped that the book is suitable for a wide class of reader, including experimental research worker and student alike, and should enable them to perform quantum-mechanical calculations and read the theoretical literature. In spite of its lack of mathematical rigour, the book should serve the theoretical student as an introduction to more advanced works.

In my approach to quantum mechanics, I have been greatly influenced by the writings of J. von Neumann and by Dr. J. S. deWet (Oxford) whose lectures first introduced me to some of these ideas, and it is a pleasure to thank him for many discussions since then. I am very grateful to Dr. Handel Davies (Oxford), with whom I discussed many aspects of this book, for a critical reading of the manuscript and for suggesting many substantial improvements, to Dr. J. M. Cassels (Liverpool), for some valuable criticisms of the text from the viewpoint of an experimentalist, and to my wife, Betty, who is responsible for many clarifications and who helped me at all stages of proof reading. I am indebted to many of my colleagues for helpful discussions arising out of my lectures, to Dr. E. Bretscher (A.E.R.E.) for encouragement to write this book and to the Director of A.E.R.E., Harwell, for permission to publish it.

Harwell, Berks
April, 1954

F. M.

I

MATHEMATICAL INTRODUCTION

THE physical concepts underlying quantum mechanics are radically different from those of classical physics. It is all the more helpful that the same mathematical methods, developed mainly by mathematicians of the nineteenth century, are used in the two theories.

In this chapter we shall introduce, and develop as far as required, the main ideas of these mathematical methods, particularly those concerning eigenvalue problems, illustrating them by a simple problem of classical physics, that of the vibrating string. We shall not introduce any new physical ideas until Chapter II.

We shall not aim at mathematical rigour or generality. What we shall do, is to present the underlying intuitive ideas of this very beautiful branch of mathematics as clearly as possible. We shall see that the mathematical formalism lends itself to a strongly geometrical interpretation which helps towards its understanding. To formulate this geometrical picture we would recall some of the basic ideas of vector analysis.

1. Vectors

Let us, to begin with, consider vectors in ordinary three-dimensional space. We shall however use a notation slightly different from the conventional one which will be more suitable for our purposes.

We introduce a system of rectangular coordinates x_1, x_2, x_3 (usually denoted by x, y, z). A point P in space is then labelled by giving its coordinates (x_1, x_2, x_3). We define a vector $\mathbf{u}$ as having components $u(1)$, $u(2)$, $u(3)$, in this coordinate system. The simplest example of a vector is the position vector $\mathbf{r}$ of the point P; its components are simply the coordinates of P: $r(1) = x_1$, *etc.* Other physical quantities which are vectors are velocity, force and acceleration. If $\mathbf{u}$ and $\mathbf{v}$ are vectors, then $c_1\mathbf{u} + c_2\mathbf{v}$ where c_1 and c_2 are real numbers, is the vector with components $c_1u(1) + c_2v(1)$, *etc.*

We define the *norm* $\mathcal{N}(\mathbf{u})$ of the vector $\mathbf{u}$ by

$$\mathcal{N}(\mathbf{u}) = [u(1)]^2 + [u(2)]^2 + [u(3)]^2$$

$$= \sum_{r=1}^{3} [u(r)]^2 \tag{1}$$

and we see that the norm is positive for any non-zero vector. The *length* or *magnitude* of $\mathbf{u}$ is defined by

$$\|\mathbf{u}\| = \sqrt{\mathcal{N}(\mathbf{u})} \tag{2}$$

where the positive square root is to be taken. For the position vector **r**,

$$\|\mathbf{r}\| = \sqrt{(x_1^2 + x_2^2 + x_3^2)}.$$

We call **u** a *normalized* or *normed* vector or a *unit* vector if it is of unit length

$$\|\mathbf{u}\| = 1. \tag{3}$$

Since $c\mathbf{u}$ (where c is a real number) is a vector parallel to **u**, of magnitude $c\|\mathbf{u}\|$, we can normalize any vector **u** by multiplying by a factor $1/\|\mathbf{u}\|$, the normed vector being $\mathbf{u}/\|\mathbf{u}\|$.

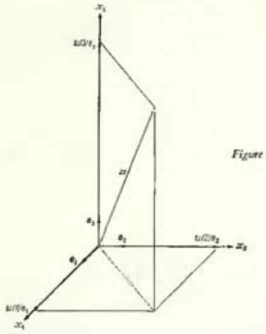

Figure 1. Decomposition of a unit vector **u** *into its components*

Given two vectors **u** and **v** (components $u(r)$ and $v(r)$, $r = 1, 2, 3$), we define their *scalar* or *inner product* by

$$\begin{aligned}(\mathbf{u}, \mathbf{v}) &= u(1)v(1) + u(2)v(2) + u(3)v(3)\\ &= \sum_{r=1}^{3} u(r)v(r).\end{aligned} \tag{4}$$

As is well known, the inner product can also be expressed in terms of the magnitudes of **u** and **v** and the angle θ between the two vectors,

$$(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos \theta. \tag{4*}$$

From equations (1) and (4) the inner product of a vector **u** with itself is its norm,

$$\mathcal{N}(\mathbf{u}) = (\mathbf{u}, \mathbf{u}). \tag{5}$$

Two vectors $\mathbf{u}$ and $\mathbf{v}$ are *normal* or *orthogonal* if $\theta=\pi/2$, *i.e.* their inner product vanishes,

$$(\mathbf{u}, \mathbf{v})=0. \tag{6}$$

If we introduce three unit vectors $\mathbf{e}_r$ $(r=1, 2, 3)$ in the directions of the coordinate axes x_r we can write any vector $\mathbf{u}$ in the form

$$\begin{aligned}\mathbf{u}&=u(1)\mathbf{e}_1+u(2)\mathbf{e}_2+u(3)\mathbf{e}_3\\&=\sum_{r=1}^{3}u(r)\mathbf{e}_r.\end{aligned} \tag{7}$$

The geometrical interpretation of this equation is shown in *Figure 1.* But there is nothing special about the three orthogonal unit vectors $\mathbf{e}_r$, which distinguishes them from any other set of three vectors $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \boldsymbol{\varphi}_3$ which are orthogonal and normed or, as we shall say, *orthonormal.* This orthonormal property can be conveniently written as

$$(\boldsymbol{\varphi}_r, \boldsymbol{\varphi}_s)=\delta_{rs} \quad (r, s=1, 2, 3) \tag{8}$$

where δ_{rs} is the Kronecker δ symbol, which is defined by

$$\begin{aligned}\delta_{rs}&=0, \quad r\neq s,\\&=1, \quad r=s.\end{aligned}$$

Introducing the set of unit vectors $\boldsymbol{\varphi}_r$ instead of the $\mathbf{e}_r$ means that we refer the components of the vectors to a different system of rectangular coordinates. For any vector $\mathbf{u}$ we can now write

$$\mathbf{u}=\sum_{r=1}^{3}c_r\boldsymbol{\varphi}_r \tag{9}$$

and the coefficients c_r in the 'expansion of $\mathbf{u}$ in terms of the $\boldsymbol{\varphi}_r$' are easily found by taking the inner product of equation (9) with $\boldsymbol{\varphi}_s$. Using (8) this gives

$$c_s=(\boldsymbol{\varphi}_s, \mathbf{u})=\sum_{r=1}^{3}\varphi_s(r)u(r). \tag{10}$$

We see that we can define a large number of different coordinate systems. But the physically significant quantities are the vectors themselves and certain functions of these, which are independent of the coordinate system used. Above, we had some examples of such quantities which are independent of the choice of coordinates: the norm of a vector, the magnitude of a vector and the inner product of two vectors.

The conditions (8) imply that we restrict ourselves to rectangular coordinate systems. If we also admit oblique Cartesian coordinates the $\boldsymbol{\varphi}_r$ need neither be normed nor orthogonal. This restriction is just convenient in practice. But any three non-coplanar vectors $\mathbf{f}^{(1)}$, $\mathbf{f}^{(2)}$, $\mathbf{f}^{(3)}$ define a coordinate system, *i.e. any* vector $\mathbf{u}$ can be written as a linear superposition

$$\mathbf{u}=c_1\mathbf{f}^{(1)}+c_2\mathbf{f}^{(2)}+c_3\mathbf{f}^{(3)}. \tag{11}$$

A set of vectors which has this property is called *complete*; we say it *spans* the vector space, *i.e.* there is no vector in the space which is orthogonal to all vectors of the set. In ordinary three-dimensional space any set of three non-coplanar vectors is complete. Two vectors $\mathbf{f}^{(1)}$ and $\mathbf{f}^{(2)}$ define a plane and all vectors of the form $c_1\mathbf{f}^{(1)}+c_2\mathbf{f}^{(2)}$ lie in this plane. We can think of this plane as a two-dimensional subspace or manifold of the three-dimensional space.

Given n vectors $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \ldots \mathbf{f}^{(n)}$, we say that they are *linearly independent* if no relation of the form

$$c_1\mathbf{f}^{(1)}+c_2\mathbf{f}^{(2)}+\ \ldots\ +c_n\mathbf{f}^{(n)}=0 \tag{12}$$

exists between them, unless $c_1=c_2=\ \ldots\ =c_n=0$. If one or more relations of the form (12) exist, with not all of the coefficients $c_1, c_2, \ldots c_n$ zero, then we say that the set $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \ldots \mathbf{f}^{(n)}$ is linearly dependent. It is obvious that in three-dimensional space the maximum number of linearly independent vectors is three.

One can clearly generalize all the above to an $\mathcal{N}$-dimensional vector space, which we shall denote by $\mathfrak{R}_{\mathcal{N}}$ (so far we considered $\mathcal{N}=3$, now $\mathcal{N}$ may be any positive integer). The formal generalizations of the concepts we introduced are quite obvious. A vector $\mathbf{u}$ of $\mathfrak{R}_{\mathcal{N}}$ has now $\mathcal{N}$ components $u(1), u(2), \ldots u(\mathcal{N})$, referred to a given coordinate system. The inner product of two vectors $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$(\mathbf{u}, \mathbf{v})=\sum_{r=1}^{\mathcal{N}}u(r)v(r), \tag{4a}$$

$$\mathcal{N}(\mathbf{u})=(\mathbf{u}, \mathbf{u})=\sum_{r=1}^{\mathcal{N}}u^2(r), \tag{5a}$$

$\mathbf{u}$ and $\mathbf{v}$ are orthogonal if $(\mathbf{u}, \mathbf{v})=0$, and $\mathbf{u}$ is a unit vector if

$$\|\mathbf{u}\|=\sqrt{\mathcal{N}(\mathbf{u})}=1. \tag{3a}$$

If we specify the coordinate system by a set of $\mathcal{N}$ orthonormal vectors $\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \ldots \boldsymbol{\varphi}_{\mathcal{N}}$,

$$(\boldsymbol{\varphi}_r, \boldsymbol{\varphi}_s)=\delta_{rs} \quad (r, s=1, 2, \ldots \mathcal{N}) \tag{8a}$$

then any vector $\mathbf{u}$ of $\mathfrak{R}_{\mathcal{N}}$ can be written as a linear superposition

$$\mathbf{u}=\sum_{r=1}^{\mathcal{N}}c_r\boldsymbol{\varphi}_r. \tag{9a}$$

A set of n vectors $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \ldots \mathbf{f}^{(n)}$ is linearly independent if the relation

$$c_1\mathbf{f}^{(1)}+c_2\mathbf{f}^{(2)}+\ \ldots\ +c_n\mathbf{f}^{(n)}=0 \tag{12a}$$

is only satisfied for $c_1=c_2=\ \ldots\ =c_n=0$. A set of n linearly independent vectors $\mathbf{f}^{(1)}, \mathbf{f}^{(2)}, \ldots \mathbf{f}^{(n)}$ span an n-dimensional space which is a subspace of $\mathfrak{R}_{\mathcal{N}}$ if $n<\mathcal{N}$ and coincides with $\mathfrak{R}_{\mathcal{N}}$ if $n=\mathcal{N}$; all vectors of the form $c_1\mathbf{f}^{(1)}+c_2\mathbf{f}^{(2)}+\ \ldots\ c_n\mathbf{f}^{(n)}$ lie in this subspace.

This last example illustrates how exceedingly helpful an intuitive geometrical interpretation of the mathematical formalism is to its clearer understanding. The mathematics can, in a sense, be interpreted as if we were still dealing with the three-dimensional case and the geometrical interpretation of orthogonality, coordinate systems, subspaces, *etc*, retains its significance. We shall frequently employ this geometrical language to present the mathematics more vividly.

By increasing the integer $\mathcal{N}$ we can let $\mathcal{N}$ tend to infinity and in this way obtain an infinitely dimensional vector space*. The purely formal mathematics and geometrical interpretation remain the same as before. A vector $\mathbf{u}$ now has infinitely many components $u(1)$, $u(2)$, $u(3)$, . . . ; its norm is

$$\mathcal{N}(\mathbf{u}) = (\mathbf{u}, \mathbf{u}) = \sum_{=1}^{\infty} u^2(r). \tag{5b}$$

This equation shows the new feature of the infinitely dimensional case. Where we previously had finite sums of $\mathcal{N}$ terms, we now have infinite series. In order that we should always obtain physically sensible results and that the mathematics should be meaningful we must impose restrictions so that all infinite series which occur are convergent. It can be shown that this is ensured if one requires all vectors $\mathbf{u}$ to have finite norms, *i.e.* for the infinite series in (5b) to be convergent. Such vectors then also have finite length. If this restriction is imposed all the mathematical expressions that occur in the theory are finite.

2. Vibrating string

Almost all of the problems of wave mechanics reduce to the solution of a differential equation with certain boundary conditions. Let us now consider a particularly simple example of such an equation, the classical problem of a vibrating string, treating and developing aspects which will be of use in all problems of this class.

We require the transverse displacement $u(x, t)$ of a string of length L, whose end points are held fixed at $(0, 0)$ and $(L, 0)$. The string is initially displaced and held at rest in a given position $u(x, t) = f(x)$ for $t < 0$, and at $t = 0$ it is released. Mathematically, this problem is expressed by the following boundary value problem (*cf* Churchill[1], p. 21):

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \tag{1}$$

$$u(0, t) = u(L, t) = 0, \quad \frac{\partial u(x, 0)}{\partial t} = 0, \tag{2}$$

$$u(x, 0) = f(x), \tag{3}$$

where c is the velocity of propagation of the wave along the string.

---

* It should be noted that the number of dimensions of this infinitely dimensional space is countable.

If we ignore equation (3) for the moment, the problem defined by equations (1) and (2) has the very important property that if u_1 and u_2 are two possible solutions, then any linear combination $c_1u_1+c_2u_2$ (c_1 and c_2 constants) is also a solution. We call problems with this property *linear* and *homogeneous*. One can try to satisfy equation (3) by taking a suitable combination of different solutions of the homogeneous problem. We try to solve equations (1) and (2) by separating the variables, *i.e.* we look for a solution of the form $u(x, t)=X(x)\ T(t)$. Equation (1) then gives

$$\frac{1}{X}\frac{d^2X}{dx^2}=\frac{1}{c^2T}\frac{d^2T}{dt^2}=\text{const.}=-\lambda. \tag{4}$$

That these expressions in X and in T are equal to a constant (which we call $-\lambda$), is due to the fact that the first is a function of x only, the second of t only. From (4) and (2) the problem for $X(x)$ becomes

$$\left.\begin{aligned} &\frac{d^2X}{dx^2}+\lambda X=0\\ &X(0)=X(L)=0. \end{aligned}\right\} \tag{5}$$

Equations (5) possess the trivial solution $X(x)\equiv 0$ for all values of the parameter λ. The question arises, whether there are special values of λ for which (5) has non-trivial solutions. We are here dealing with a very simple but typical *eigenvalue problem*, *i.e.* we have a differential equation containing a parameter λ, and we require solutions satisfying certain boundary conditions. If there are special values of λ for which non-trivial solutions exist, we call these *eigenvalues*, and the corresponding solutions *eigensolutions* or *eigenfunctions*. The general solution of the differential equation in (5) is

$$X(x)=A\sin\sqrt{\lambda}x+B\cos\sqrt{\lambda}x,$$

and the boundary conditions determine A and B.

$$X(0)=0 \quad \text{implies} \quad B=0,$$
$$X(L)=0 \quad \text{implies} \quad A\sin(\sqrt{\lambda}L)=0;$$

$A=0$ is the trivial and not very useful solution $u\equiv 0$; hence we must have $\sin(\sqrt{\lambda}L)=0$, *i.e.* $(\sqrt{\lambda}L)=n\pi$ $(n=1, 2, \ldots)$.

Thus we obtain a series of eigenfunctions

$$X_n(x)=\sin\frac{n\pi}{L}x \quad (n=1, 2, \ldots) \tag{6}$$

and corresponding eigenvalues

$$\lambda_n=\left(\frac{n\pi}{L}\right)^2 \quad (n=1, 2, \ldots). \tag{7}$$

To solve the corresponding problem for $T(t)$ we are no longer at liberty to choose the value of λ arbitrarily, but must use one of the values λ_n defined in (7). Equations (4) and (2) now give

$$\left.\begin{aligned} \frac{1}{c^2 T}\frac{d^2 T}{dt^2} &= -\lambda_n \\ \frac{dT}{dt} &= 0, \text{ at } t=0. \end{aligned}\right\}$$

In the general solution

$$T(t) = C\cos(\sqrt{\lambda_n}ct) + D\sin(\sqrt{\lambda_n}ct)$$

the boundary condition again leads to $D=0$.

The general solution of our boundary value problem (1), (2) is hence a linear superposition of the solutions of the form

$$\sin\left(\frac{n\pi}{L}x\right)\cos\left(\frac{n\pi}{L}ct\right)$$

i.e.

$$u(x,t) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{L}x\right)\cos\left(\frac{n\pi}{L}ct\right) \tag{8}$$

where the A_n are as yet undetermined constants. To find these A_n, we use (3), *i.e.* for $t=0$, equation (8) is to reduce to

$$f(x) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi}{L}x\right). \tag{9}$$

The infinite series on the right-hand side of (9) is called a *Fourier* series (or a Fourier sine series, as it only involves sines and not cosines). The important question now is the following: given $f(x)$, can we always find constants A_n $(n=1, 2, \ldots)$, so that (9) is satisfied? Using the easily proved relations

$$\int_0^L \sin\left(\frac{m\pi}{L}x\right)\sin\left(\frac{n\pi}{L}x\right)dx = \frac{L}{2}\delta_{mn}, \quad (m, n = 1, 2, \ldots) \tag{10}$$

we can at once derive coefficients A_n which satisfy (9). We multiply (9) by $\sin\left(\frac{m\pi}{L}x\right)$ and integrate with respect to x from 0 to L. On account of the δ_{mn} in (10) only the term with $n=m$ of the infinite series in (9) gives a non-vanishing contribution, namely

$$A_m = \frac{2}{L}\int_0^L f(x)\sin\left(\frac{m\pi}{L}x\right)dx \quad (m=1, 2, \ldots). \tag{11}$$

These coefficients A_m complete the solution of our inhomogeneous boundary value problem, equations (1) to (3), the solution being

$$u(x,t) = \sum_{n=1}^{\infty}\left[\frac{2}{L}\int_0^L f(\xi)\sin\left(\frac{n\pi}{L}\xi\right)d\xi\right]\sin\left(\frac{n\pi}{L}x\right)\cos\left(\frac{n\pi}{L}ct\right). \tag{12}$$

The derivation of this solution has been purely formal, *i.e.* we said nothing about the convergence of the infinite series in (9), nor did we justify the term-wise integration of this series to derive the A_n*. The rigorous justification that, with very weak restrictions on the function $f(x)$, the term-wise integration of this series is permissible and that the infinite series in (9) converges to the prescribed value $f(x)$ forms the central problem of the theory of Fourier series. We shall not carry out this proof. From (9) we see that $f(x)$ clearly must satisfy these conditions

(*i*) $f(x) = -f(-x)$, *i.e.* $f(x)$ is an *odd* function,
(*ii*) $f(x+2L) = f(x)$, *i.e.* $f(x)$ is periodic, with period $2L$.

In addition $f(x)$ must satisfy certain continuity conditions†. If $f(x)$ satisfies these conditions, then its Fourier series converges at every point x‡ and (9) is satisfied. In our problem of the vibrating string, these conditions are always satisfied. Since $f(x)$ is defined in the interval $(0, L)$ only, it can always be continued in $(-L, 0)$ in such a way as to satisfy conditions (*i*) and (*ii*).

By a modification of the definition of convergence, one can enlarge the class of functions which can be expanded in a Fourier sine series to consist of all odd functions of period $2L$ for which the integral

$$\int_0^L \{f(x)\}^2 dx$$

is finite. In quantum mechanics it is this modified convergence which is relevant and, therefore, we shall always consider this wider class of functions hereafter.

3. Mathematical generalizations

If $f(x)$ satisfies the conditions stated at the end of section 2, it can be expanded in a Fourier sine series

$$f(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x) \tag{1}$$

where

$$\varphi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi}{L}x\right) \quad (n = 1, 2, \ldots); \tag{2}$$

---

* For a finite number of terms the integral of a sum is always equal to the sum of the integrals; for an infinite series this is not necessarily the case and needs a special investigation to justify it rigorously.

† A continuity condition sufficiently general to cover all applications, is the following: (x) is sectionally smooth in the interval $0 < x < L$, *i.e.* $f(x)$ and $f'(x)$ are continuous in this interval except for a finite number of jump discontinuities. At such a point of discontinuity, $x = x_0$ say, we define $f(x_0)$ as

$$f(x_0) = \tfrac{1}{2}\{f(x_0+0) + f(x_0-0)\}$$

where $f(x_0+0)$ and $f(x_0-0)$ are the values of $f(x)$ as the point $x = x_0$ is approached from the right and from the left respectively.

‡ At a point of discontinuity, $x = x_0$, it converges to the arithmetic mean $f(x_0)$ defined above.

we include the factor $\surd(2/L)$ in the definition of φ_n so that equation (2.10) should take the form

$$\int_0^L \varphi_m(x)\varphi_n(x)dx=\delta_{mn} \quad (m, n=1, 2, \dots). \tag{3}$$

The coefficients c_n in (1) are now given by

$$c_n=\int_0^L \varphi_n(x)f(x)dx \quad (n=1, 2, \dots). \tag{4}$$

Equations (1) to (4) show a close correspondence between vector spaces, treated in section 1, and Fourier series expansions. To bring this correspondence out, we introduce the terminology of vectors into the theory of Fourier series.

We define the norm $\mathcal{N}\{f(x)\}$ of a function $f(x)$ by

$$\mathcal{N}\{f(x)\}=\int_0^L f^2(x)dx \quad [cf\ (1.1)]. \tag{5}$$

The length or magnitude of $f(x)$ is defined by

$$\|f(x)\|=\surd\mathcal{N}\{f(x)\} \quad [cf\ (1.2)] \tag{6}$$

and $f(x)$ is normalized or normed if

$$\|f(x)\|=1 \quad [cf\ (1.3)]. \tag{7}$$

Given a function $f(x)$, we can form a normed function from it:

$$\frac{1}{\|f(x)\|}f(x).$$

Given two functions $f(x)$ and $g(x)$, we define their scalar or inner product

$$(g(x),f(x))=\int_0^L g(x)f(x)dx \quad [cf\ (1.4)]. \tag{8}$$

In particular, the norm of $f(x)$ is the inner product of $f(x)$ with itself

$$\mathcal{N}\{f(x)\}=(f(x),f(x)) \quad [cf\ (1.5)]. \tag{9}$$

Two functions $f(x)$ and $g(x)$ are orthogonal if their inner product vanishes,

$$(g(x),f(x))=0 \quad [cf\ (1.6)]. \tag{10}$$

We see that the functions $\varphi_n(x)$, equations (2), form a normed orthogonal sequence, or as we shall say, an orthonormal sequence,

$$(\varphi_m(x), \varphi_n(x))=\delta_{mn} \quad [cf\ (1.8)] \tag{3}$$

and any function satisfying certain conditions, stated above, can be written as a linear superposition of these $\varphi_n(x)$,

$$f(x)=\sum_{n=1}^{\infty} c_n\varphi_n(x) \quad [cf\ (1.9)]. \tag{1}$$

3. MATHEMATICAL GENERALIZATIONS

The close analogy which seems to exist between vector analysis and Fourier analysis suggests the consistent use of 'vector-language' in the Fourier analysis. Nothing new is obtained this way which is not already contained in the mathematics, but one obtains a strongly geometrical language which gives one a good feeling for the mathematics which is most useful.

We shall say that the set $\varphi_1(x), \varphi_2(x), \ldots$ form a coordinate system, *i.e.* a set of orthogonal unit vectors defining an infinitely dimensional *function space*. The vectors of this space are the functions which can be expanded in a Fourier sine series of period $2L$, equation (1). The function space which we have here defined, has structurally the same properties as the infinitely dimensional vector space we introduced in section 1.

We have here defined the function space by specifying a coordinate system $\varphi_1, \varphi_2, \ldots,$ and defining the space to consist of all linear combinations

$$\sum_{n=1}^{\infty} c_n \varphi_n(x)$$

of finite length. (Just as in section 1, we impose this restriction to exclude physically and mathematically meaningless results.) Using equations (8), (9) and (3) we obtain

$$\begin{aligned} \mathcal{N}\left\{\sum_{n=1}^{\infty} c_n \varphi_n(x)\right\} &= \left(\sum_{m=1}^{\infty} c_m \varphi_m(x), \sum_{n=1}^{\infty} c_n \varphi_n(x)\right) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m c_n (\varphi_m(x), \varphi_n(x)) \\ &= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} c_m c_n \delta_{mn} \\ &= \sum_{n=1}^{\infty} c_n^2, \end{aligned}$$

i.e., as in section 1, we require $\sum_{n=1}^{\infty} c_n^2$ to be finite. Now just as in the case of vector spaces, the physically significant quantities are the functions, not the way in which they are expanded, and we can find many other coordinate systems spanning the function space, *i.e.* any function of it, $\sum c_n \varphi_n(x)$, can be written as a linear combination of the new normed orthogonal set of functions. We can express this differently by saying that there is no vector of the space which is orthogonal to all vectors of the set. A set with this property is called *complete*. Just as for vector spaces, a set of functions must be complete to form a coordinate system, but they need not be orthogonal or normed. But in practice a complete normalized orthogonal set of functions is usually most convenient.

We again define a set of functions $f_1(x), f_2(x), \dots f_n(x)$ to be *linearly independent* if the relation

$$c_1 f_1(x) + c_2 f_2(x) + \dots + c_n f_n(x) = 0 \quad [cf\ (1.12)] \qquad (11)$$

is only satisfied for $c_1 = c_2 = \dots = c_n = 0$. For example, if $f_1, f_2, \dots f_n$ are orthogonal they are clearly linearly independent. For taking the inner product of (11) with f_m $(m = 1, 2, \dots n)$ gives at once, on account of $(f_r, f_s) = \delta_{rs}$, that $c_m = 0$, $(m = 1, 2, \dots n)$.

Instead of defining the function space in terms of a complete set of functions, which serve as a coordinate system, we can define it by specifying the conditions which any function of the space must satisfy. In accordance with the discussion at the end of section 2 we could, in our case, define the function space to consist of all odd functions of period $2L$ which possess a finite norm. This at once suggests that we can not only span the same function space by different coordinate systems but that we can define different function spaces. Thus by considering different values of L in the above example, we obtain different function spaces. Or we could drop the restriction to odd functions, the resulting function space being spanned by both sine and cosine functions (*cf* Exercises I.1–I.4). In the general case one considers a certain fundamental domain or region corresponding to the physical system one wants to describe. In the above case of a string of length L it was the domain $0 \leqslant x \leqslant L$. In general it will be some interval $a \leqslant x \leqslant b$ which may be the whole of the x-axis $-\infty < x < \infty$, or it may be a three-dimensional region corresponding to the space occupied by the physical system under investigation. All integrals occurring in inner products are to be taken over the fundamental domain and we shall not in general indicate the limits of integration. Corresponding to a given fundamental domain we shall consider the function space consisting of all functions f which have a finite norm and which satisfy certain boundary conditions at the boundary of the fundamental domain, determined by the physical problem (*e.g.* in the example of the vibrating string the conditions that its ends were fixed imposed the boundary conditions $f(0) = f(L) = 0$).

There is, however, one other generalization which must be made. In quantum mechanics we often deal with *complex* functions and our definitions of inner product, *etc* must then be modified. The corresponding function space is called a *Hilbert space.* We define the inner product of $f(x)$ and $g(x)$ as

$$(g(x), f(x)) = \int \overline{g(x)} f(x) dx \qquad (8a)$$

where $\overline{g(x)}$ is the complex conjugate of $g(x)$. (Note that x itself is always real but $g(x)$ may be a complex function of x, *e.g.* e^{ix}). From

(8a) it follows that

$$(g(x), f(x)) = \overline{(f(x), g(x))}, \tag{12}$$

i.e. in general the inner product is now a complex number and the order of the factors cannot be interchanged. The norm, now defined by

$$\mathcal{N}\{f(x)\} = (f(x), f(x)) = \int \overline{f(x)} f(x) dx = \int |f(x)|^2 dx \tag{5a}$$

is always real and

$$\|f(x)\| = \sqrt{\mathcal{N}\{f(x)\}} \tag{6a}$$

is real and positive (and vanishes if and only if $f(x) \equiv 0$). In an expansion

$$f(x) = \sum c_n \varphi_n(x), \tag{13}$$

of a function $f(x)$ in terms of some complete normed orthogonal set of functions, $\varphi_1, \varphi_2, \ldots$, the expansion coefficients c_n are now complex numbers

$$c_n = (\varphi_n(x), f(x)), \tag{14}$$

and from (8a) it follows that for two complex numbers a and b

$$(bg(x), af(x)) = \bar{b}a(g(x), f(x)). \tag{15}$$

As an example of a complete orthonormal set of complex functions, we have the complex exponential functions

$$\varphi_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, \quad n = 0, \pm 1, \pm 2, \ldots, \tag{16}$$

which span the space of all functions of period 2π (with finite norm). The fundamental domain is $0 < x < 2\pi$, and one has

$$(\varphi_m, \varphi_n) = \frac{1}{2\pi} \int_0^{2\pi} e^{i(n-m)x} dx = \begin{cases} 0, & n \neq m \\ 1, & (n = m). \end{cases}$$

In section 2 we considered one of the simplest possible eigenvalue problems and saw that the set of all eigenfunctions form a complete normed orthogonal sequence which can serve as a coordinate system of the function space. In quantum mechanics more general eigenvalue problems occur which are essentially of the same type, *i.e.* they consist of a differential equation containing a parameter λ. One requires solutions satisfying the differential equation in a prescribed fundamental region and satisfying certain boundary conditions on its boundary. Such solutions do not exist for all values of λ, but only for certain values $\lambda_1, \lambda_2, \ldots$, called the eigenvalues. The corresponding solutions $\varphi_1, \varphi_2, \ldots$, the eigenfunctions, form a complete orthonormal set spanning the function space consisting of all functions which have a finite norm over the fundamental domain. The rigorous proof of the completeness of the set of eigenfunctions, *i.e.* the possibility of expanding an arbitrary function satisfying the above conditions in terms of these eigenfunctions, forms a very difficult branch of pure mathematics. The proof can be carried out in all practical cases. Accordingly, we shall always assume the completeness of the set of eigenfunctions in all eigenvalue problems we encounter.

A new feature arising in these more general eigenvalue problems is the existence of *degenerate* eigenvalues. We call an eigenvalue λ_0 N-fold degenerate if, corresponding to it, there exist N linearly independent eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_N$. Any linear combination

$$c_1\varphi_1 + c_2\varphi_2 + \ \ldots \ + c_N\varphi_N \tag{17}$$

then is also an eigenfunction corresponding to the same eigenvalue λ_0. Conversely, any eigenfunction corresponding to the eigenvalue λ_0 is a linear combination of the set $\varphi_1, \varphi_2, \ldots \varphi_N$. We can express the above definition of the degeneracy N slightly differently by saying that N is the least number of eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_N$, so that all eigenfunctions corresponding to the eigenvalue λ_0 can be written in the form (17). We can say that the set $\varphi_1, \varphi_2, \ldots \varphi_N$, spans an N-dimensional manifold or subspace; all functions of this subspace are eigenfunctions corresponding to λ_0. We can of course span this subspace in a large number of different ways. In the set $\varphi_1, \varphi_2, \ldots \varphi_N$, we could for example replace φ_1 and φ_2 by any two linear combinations

$$\psi_1 = a_1\varphi_1 + a_2\varphi_2, \quad \psi_2 = b_1\varphi_1 + b_2\varphi_2,$$

with $a_1/a_2 \neq b_1/b_2$. Clearly, the eigenfunctions corresponding to λ_0 are not all mutually orthogonal; in fact, they form a continuous subspace. But we can always span the subspace by a set of normed orthogonal eigenfunctions. This is intuitively obvious and we shall only illustrate it for $N=2$. Suppose φ_1 and φ_2 are two independent eigenvectors belonging to the eigenvalue λ_0. Then any vector lying in the plane defined by φ_1 and φ_2, *i.e.* of the form $c_1\varphi_1 + c_2\varphi_2$, is also an eigenvector of λ_0. We can replace the oblique coordinates φ_1, φ_2, by a rectangular coordinate system ψ_1, ψ_2, but this normed orthogonal system is of course not uniquely defined. Suppose we take

$$\psi_1 = \frac{1}{\|\varphi_1\|}\varphi_1.$$

If we write φ_2 in the form of a multiple of ψ_1 plus a vector normal to ψ_1,

$$\varphi_2 = a\psi_1 + \chi_2$$

then a is determined by taking the inner product with ψ_1,

$$a = (\psi_1, \varphi_2).$$

Hence

$$\chi_2 = \varphi_2 - \psi_1(\psi_1, \varphi_2)$$

is orthogonal to ψ_1, $(\psi_1, \chi_2) = 0$, as is easily verified. To obtain ψ_2 we must finally normalize χ_2,

$$\psi_2 = \frac{1}{\|\chi_2\|}\chi_2.$$

The geometrical interpretation of this orthogonalization process for the case of real functions is illustrated in *Figure 2*. This method, known as the Schmidt orthogonalization process, can clearly be generalized.

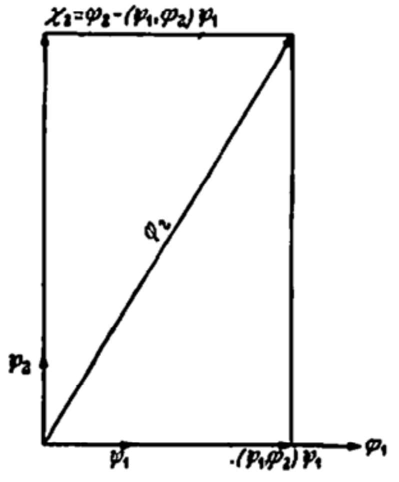
$\chi_2=\varphi_2-(\psi_1,\varphi_2)\psi_1$
φ_2
ψ_2
ψ_1
$(\psi_1\varphi_2)\psi_1$
φ_1

Figure 2. The Schmidt orthogonalization process (ψ_1, φ_2) ψ_1 is the projection of φ_2 on the ψ_1-axis

4. Continuous spectra

Another new feature of general eigenvalue problems is the occurrence of continuous spectra of eigenvalues. It may happen that instead of a discrete set of eigenvalues $\lambda_1, \lambda_2, \ldots$, (we shall also call this a *discrete spectrum* of eigenvalues), the eigenvalues may fill a continuous range (we call this a *continuous spectrum* of eigenvalues). An arbitrary function is then not represented as an infinite series but as an integral. Such continuous spectra occur in quantum mechanics only in problems with infinite fundamental domains. We can see how they come about by returning to the problem of a vibrating string treated in section 2. There we obtained the eigenvalues

$$\lambda_n = \left(\frac{n\pi}{L}\right)^2, \quad (n = 1, 2, \ldots). \tag{2.7}$$

If we increase the fundamental domain, *i.e.* we increase L, the eigenvalues move closer and closer together. In the limit as $L \to \infty$, they form a continuum. Let us investigate this behaviour more closely in two special examples.

Example 1. Consider the eigenvalue problem

$$\frac{d^2X}{dx^2}+\lambda X=0, \quad -\infty<x<\infty$$

with the boundary conditions

$$X(x) \text{ finite, as } \quad x\rightarrow\pm\infty.$$

All real positive λ are two-fold degenerate eigenvalues, the corresponding eigenfunctions being

$$\varphi(x)=\exp\{\pm i\sqrt{\lambda}x\}.$$

These functions certainly satisfy the eigenvalue problem but they cannot be normalized

$$(\varphi,\varphi)=\int_{-\infty}^{\infty}|\varphi(x)|^2dx=\int_{-\infty}^{\infty}dx.$$

An arbitrary function $f(x)$ is now not represented by a Fourier series but by a Fourier integral

$$f(x)=\int_{-\infty}^{\infty}c(\nu)e^{i\nu x}d\nu.$$

In the theory of Fourier integrals one shows that such a representation is always possible provided that

$$\mathcal{N}\{f(x)\}=\int_{-\infty}^{\infty}|f(x)|^2dx$$

is finite and the expansion coefficient $c(\nu)$ is then given by

$$c(\nu)=\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x)e^{-i\nu x}dx.$$

Example 2. Consider the eigenvalue problem

$$\frac{1}{r}\frac{d^2}{dr^2}(ru)+k^2u=0 \tag{1}$$

with the boundary conditions

$$u(0) \text{ finite,} \quad u(a)=0.$$

This is the problem of standing radial waves $u(r)$ inside a sphere of radius a. If we put $\psi(r)=ru(r)$, this boundary value problem takes the form

$$\frac{d^2}{dr^2}\psi+k^2\psi=0 \tag{2}$$

with the boundary conditions

$$\psi(0)=\psi(a)=0, \tag{3}$$

i.e. we have reduced the original boundary value problem to that of the vibrating string treated in section 2. We hence obtain the eigensolutions

$$u_n(r)=\frac{1}{r}\psi_n=A_n\frac{1}{r}\sin\frac{n\pi}{a}r\quad(n=1,2,\ldots),\tag{4}$$

the eigenvalues being

$$k_n=\frac{n\pi}{a}\quad(n=1,2,\ldots).\tag{5}$$

We next determine A_n so that $u_n(r)$ is normalized. We require

$$\begin{aligned}1=\mathcal{N}\{u_n(r)\}&=\int_0^a|u_n(r)|^2 4\pi r^2dr\\&=4\pi|A_n|^2\int_0^a\sin^2\left(\frac{n\pi}{a}r\right)dr\\&=2\pi a|A_n|^2,\end{aligned}$$

i.e. u_n is normed if we take A_n as real, independent of n and equal to

$$A=\frac{1}{\sqrt{2\pi a}}.\tag{6}$$

If we consider what happens as we let $a\to\infty$, we find, since

$$k_{n+1}-k_n=\frac{\pi}{a}$$

that

$$\lim_{a\to\infty}(k_{n+1}-k_n)=0,$$

and from (6) that the eigenfunctions u_n cannot be normalized: $a\to\infty$ implies $A\to 0$, *i.e.* $u_n\to 0$. Thus the functions

$$u(r)=\frac{1}{r}\sin kr$$

are eigenfunctions of our eigenvalue problem (2)–(3) (with $a\to\infty$) for all values k; but they cannot be normalized.

The impossibility of normalizing the eigenfunctions in the above examples is typical of eigenvalue problems with continuous spectra: the integrals defining the norm are divergent. This is due to the fact that the eigenfunctions do not tend to zero sufficiently rapidly at large distances; they are oscillatory instead of exponentially decreasing.

In addition to the completely discrete and completely continuous spectra so far considered, eigenvalue problems with mixed spectra occur in quantum mechanics. These are usually of the form: for $\lambda<\lambda_0$ there exist a (finite or infinite) number of discrete eigenvalues; for $\lambda>\lambda_0$ the spectrum is continuous, *i.e.* all $\lambda>\lambda_0$ are eigenvalues.

4. CONTINUOUS SPECTRA

Let us consider the case of a completely continuous spectrum in greater detail*. Consider the fundamental domain $-\infty < x < \infty$. Suppose we have a continuous set of eigenfunctions $\varphi_\mu(x)$ and corresponding eigenvalues λ_μ and assume μ has the continuous range $-\infty < \mu < \infty$. Then,

$$(\varphi_\mu, \varphi_\mu) = \int_{-\infty}^{\infty} |\varphi_\mu(x)|^2 dx \tag{7}$$

is divergent, but any function $f(x)$ for which

$$\mathcal{N}\{f(x)\} = \int_{-\infty}^{\infty} |f(x)|^2 dx \tag{8}$$

is finite can be written as an integral

$$f(x) = \int_{-\infty}^{\infty} a(\mu)\varphi_\mu(x) d\mu, \tag{9}$$

with

$$a(\mu) = (\varphi_\mu, f) = \int_{-\infty}^{\infty} \overline{\varphi_\mu(x)} f(x) dx, \tag{10}$$

analogously to (3.13) and (3.14). Equations (9) and (10) imply a certain normalization of the $\varphi_\mu(x)$. (We shall say they are normed per unit μ.)

Equations (9) and (10) are all that one requires in practice. The physically permitted functions can always be normed, *i.e.* they correspond to the $f(x)$ and not the $\varphi_\nu(x)$. They can hence always be expanded in the above form. But we can try to take the analogy with the discrete case one stage further and obtain equations analogous to (3.3)

$$(\varphi_m(x), \varphi_n(x)) = \delta_{mn}. \tag{3.3}$$

To do this, we substitute (9) into (10)

$$a(\mu) = \int_{-\infty}^{\infty} \overline{\varphi_\mu(x)} dx \int_{-\infty}^{\infty} a(\nu)\varphi_\nu(x) d\nu, \tag{11}$$

and change the order of integration, obtaining

$$a(\mu) = \int_{-\infty}^{\infty} a(\nu) d\nu \int_{-\infty}^{\infty} \overline{\varphi_\mu(x)} \varphi_\nu(x) dx. \tag{12}$$

The first integral is a function of μ and ν which we shall call $\delta(\nu - \mu)$ [as we shall presently see that it is a function only of the difference $(\nu - \mu)$ and not of ν and μ separately],

$$\delta(\nu - \mu) = \int_{-\infty}^{\infty} \overline{\varphi_\mu(x)} \varphi_\nu(x) dx, \tag{13}$$

---

* The treatment of continuous spectra that follows will seldom be needed in this book and may be omitted. It is however essential to a large part of the literature, particularly on scattering problems.

so that (12) becomes

$$a(\mu)=\int_{-\infty}^{\infty}a(\nu)\delta(\nu-\mu)d\nu. \tag{14}$$

In (14), $a(\nu)$ is now an arbitrary function. So that (14) should generally hold for any function $a(\nu)$ the function $\delta(\nu-\mu)$ must have these properties:

$$\delta(\nu-\mu)=0, \quad \text{if } \nu\neq\mu; \tag{15a}$$

but for $\nu=\mu$, $\delta(\nu-\mu)$ becomes infinite in such a way that for any interval (ν_1, ν_2) enclosing the point $\nu=\mu$

$$\int_{\nu_1}^{\nu_2}\delta(\nu-\mu)d\nu=1. \tag{15b}$$

The function $\delta(\nu-\mu)$ with the highly singular properties defined in equations (15a) and (15b) is called the *Dirac delta-function*. Clearly no ordinary mathematical function with these properties exists. That we have ended up with such an improper function is due to the fact that the change of order of integration to pass from equation (11) to equation (12) is not permissible, and equation (12) is meaningless. Whereas the first integral in (12), namely

$$\delta(\nu-\mu)=\int_{-\infty}^{\infty}\overline{\varphi_\mu(x)}\varphi_\nu(x)dx,$$

diverges for $\nu=\mu$, due to the fact that the $\varphi_\nu(x)$ do not tend to zero sufficiently rapidly as $|x|\to\infty$, the first integral in (11), namely

$$\int_{-\infty}^{\infty}a(\nu)\varphi_\nu(x)d\nu,$$

converges and tends to zero as $|x|\to\infty$ due to interference between oscillatory functions $\varphi_\nu(x)$ with different values of ν.

In spite of these mathematical objections to the Dirac δ-function, it is a most convenient function to use symbolically. For in applications the δ-function always occurs under an integral sign. Carrying out this integration, using the formal properties of the δ-function, is really equivalent to inverting the order of integration once more, thus getting back to a mathematically correct expression. Thus, using (15b), we have

$$\int_{-\infty}^{\infty}g(\mu)\delta(\mu-\mu_0)d\mu=g(\mu_0),$$

but substituting (13) for the δ-function and changing the order of integration we obtain the mathematically correct expression

$$\begin{aligned}\int_{-\infty}^{\infty}g(\mu)\delta(\mu-\mu_0)d\mu&=\int_{-\infty}^{\infty}g(\mu)d\mu\int_{-\infty}^{\infty}\overline{\varphi_{\mu_0}(x)}\varphi_\mu(x)dx\\&=\int_{-\infty}^{\infty}\overline{\varphi_{\mu_0}(x)}dx\int_{-\infty}^{\infty}g(\mu)\varphi_\mu(x)d\mu,\end{aligned}$$

and comparing this expression with equations (9) and (10) [in (9) and (10) replace $a(\mu)$ by $g(\mu)$], we see at once that this double integral is equal to $g(\mu_0)$.

Finally we will state a few properties of the Dirac δ-function*.

(*a*) By choosing different possible sets of functions $\varphi_\nu(x)$, we obtain different possible representations of the δ-function from equation (13),

$$\delta(\nu-\mu)=\int_{-\infty}^{\infty}\overline{\varphi_\mu(x)}\varphi_\nu(x)dx. \tag{13}$$

We obtain a particularly useful representation if for the $\varphi_\nu(x)$ we choose the eigenfunctions of Example 1 above:

$$\varphi_\nu(x)=\frac{1}{\sqrt{2\pi}}e^{i\nu x}. \tag{16}$$

The factor $(2\pi)^{-\frac{1}{2}}$ has to be included in this definition of $\varphi_\nu(x)$ so that equations (9) and (10), from which (13) was derived, should hold. With this choice of $\varphi_\nu(x)$, equations (9) and (10) are a statement of the Fourier integral theorem. Substituting (16) into (13) we obtain an integral representation of the δ-function

$$\delta(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ix\xi}d\xi. \tag{17}$$

(*b*) From the implicit definition of $\delta(x)$, equations (15a–b), or from the explicit definition, equation (17), we can easily verify the following most frequently required properties of the δ-function:

(*i*) If $a_1<a_2$,

$$\int_{a_1}^{a_2}f(x)\delta(x-x_0)dx=\begin{cases}f(x_0), \text{ if } a_1<x_0<a_2\\ 0, \text{ if } x_0<a_1 \text{ or } x_0>a_2.\end{cases}$$

(*ii*) $\delta(-x)=\delta(x)$.

(*iii*) $\delta(\lambda x)=\frac{1}{\lambda}\delta(x) \quad (\lambda>0)$.

(*iv*) $x\delta'(x)=-\delta(x)$.

(*c*) If $\mathbf{r}$ is the vector with components (x, y, z) we define the three-dimensional δ-function

$$\delta(\mathbf{r})=\delta(x)\delta(y)\delta(z). \tag{18}$$

$\delta(\mathbf{r})$ has similar properties to $\delta(x)$; in particular

$$\int f(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}_0)d\tau=\begin{cases}f(\mathbf{r}_0), & \text{if } \mathbf{r}_0 \text{ lies inside the domain of integration,}\\ 0, & \text{if } \mathbf{r}_0 \text{ lies outside} \quad ,, \quad ,, \quad ,, \quad ,\end{cases}$$

$d\tau$ being an element of volume.

---

* A more detailed treatment of the Dirac delta function will be found in the books by Schiff[12] and Dirac[15], mentioned in the bibliography at the end of this book.

EXERCISES I

I.1. Show that for the fundamental domain $-L<x<L$, the set of functions

$$\psi_0=\frac{1}{\sqrt{2L}},\quad \psi_n=\frac{1}{\sqrt{L}}\cos\frac{n\pi}{L}x,\quad \varphi_n=\frac{1}{\sqrt{L}}\sin\frac{n\pi}{L}x,\quad (n=1,2,\ldots) \tag{I.1}$$

form a normed orthogonal set.

I.2. Given that the function $f(x)$, defined in the interval $-L<x<L$, can be expanded in a Fourier series, *i.e.*

$$f(x)=a_0\psi_0+\sum_{n=1}^{\infty}(a_n\psi_n+b_n\varphi_n), \tag{I.2}$$

where φ_n and ψ_n are defined as in Exercise I.1, what are the Fourier coefficients a_n, b_n?

Show that if $L=\pi$, the function space spanned by the ψ_n, φ_n is the same as that spanned by the complex exponential functions, equation (3.16),

$$\chi_n(x)=\frac{1}{\sqrt{2\pi}}e^{inx},\quad (n=0,\ \pm1,\ \pm2,\ldots).$$

I.3. Show that if $f(x)$ is an odd or even function respectively, the Fourier series of Exercise I.2 reduces to a Fourier sine or cosine series respectively.

I.4. Obtain the Fourier series of the function $f(x)$ defined as follows

$$\begin{aligned}
&f(x)=0, && -\pi<x<0\\
&f(x)=x, && 0<x<\pi\\
&f(\pi)=\pi/2, &&\\
&f(x+2\pi)=f(x). && \text{for } |x|>\pi.
\end{aligned}$$

I.5. Consider the function space of all functions $f(x)$ with finite norm over the fundamental domain $-1<x<1$. Use the Schmidt orthogonalization process to span the three-dimensional manifold, defined by the set of functions

$$\varphi_0(x)=1,\quad \varphi_1(x)=x,\quad \varphi_2(x)=x^2,$$

by a normed orthogonal sequence ψ_0, ψ_1, ψ_2, with ψ_0 a multiple of φ_0.

I.6. The Fourier integral theorem can be stated as follows. If $f(x)$ is a function for which $\int_{-\infty}^{\infty}|f(x)|^2dx$ is finite, then

$$f(x)=\int_{-\infty}^{\infty}c(k)e^{ikx}dk \tag{I.6a}$$

with

$$c(k)=\frac{1}{2\pi}\int_{-\infty}^{\infty}f(x)e^{-ikx}dx. \tag{I.6b}$$

(The function $c(k)$ is known as the Fourier transform of $f(x)$.) If

$$f(x)=\left\{\begin{array}{ll} e^{ik_0x}, & \text{for } |x|<L\\ 0, & \text{for } |x|>L\end{array}\right\} \tag{I.6c}$$

find $c(k)$; interpret equations (I.6a–c) and the result for $c(k)$ in terms of a superposition of waves.

I.7. **Prove Schwarz's inequality:** if φ and ψ are functions with finite norms, then

$$|(\varphi, \psi)| \leq \|\varphi\| \, \|\psi\|. \tag{I.7}$$

The equality sign holds if and only if ψ is proportional to φ, *i.e.* $\psi = a\varphi$ where a is a complex number.

(Hint: consider $(f, f) \geqslant 0$ where $f = (\varphi, \psi)\varphi - (\varphi, \varphi)\psi$.)

I.8. **Derive and solve** the following eigenvalue problem which occurs in the theory of a vibrating square membrane whose sides, of length L, are kept fixed:

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \lambda w = 0, \tag{I.8a}$$

$$\left.\begin{aligned} w(0, y) = w(L, y) = 0, \quad (0 < y < L) \\ w(x, 0) = w(x, L) = 0, \quad (0 < x < L). \end{aligned}\right\} \tag{I.8b}$$

II

WAVE-MECHANICAL CONCEPTS

At the end of the last century, physical theory, classical physics as it is now called, seemed to have reached a high degree of perfection. There existed two essentially different modes of description of natural phenomena, either in terms of a wave picture or a particle picture. The wave picture was eminently suitable for describing interference, diffraction and polarization phenomena, whereas the particle picture had failed to account for these. Consequently optical and electromagnetic processes in general were described in terms of waves or fields. Atoms and atomic constituents such as electrons, on the other hand, were described as particles, *i.e.* moving in definite paths completely determined according to Newtonian mechanics by the forces they experienced.

Experiments carried out since the beginning of this century show that, on the atomic as distinct from the macroscopic scale, matter and light display a dualistic nature. One can devise experiments in which electrons display wave properties as in electron diffraction, and others in which light displays particle properties as in the photoelectric effect.

Quantum mechanics attempts to resolve the apparent paradox of this wave-particle dualism by a new formulation of the laws of nature in which the particle and wave aspects are not mutually exclusive but complementary; they enter the theory, so to speak, on an equal footing. The two interpretations are related by Planck's constant h which in macroscopic units (*e.g.* in c.g.s. units) is exceedingly small. Hence, in general, we may put $h=0$ in the equations of quantum theory. This just leads back to classical physics: for electrons it results in a particle picture, for light in a wave picture. In the atomic domain we are dealing with quantities themselves of the same order of magnitude as Planck's constant, and severe deviations from classical physics occur. An electron should now be thought of as a 'wave-particle', displaying the typical wave or particle properties of classical physics according to the experiment performed.

We shall not treat the general formulation of quantum mechanics and its interpretation till a later stage (Chapter IV). In this chapter we want to introduce the wave equation which forms the basis of wave mechanics, the form of quantum mechanics developed by Schrödinger in 1926, and discuss its properties and interpretation. There are various ways of introducing quantum mechanics. Rather than to try to derive quantum mechanics from a detailed analysis of experiments, it seems preferable to state its concepts and laws in a simple plausible manner and to take for its ultimate justification the large field where quantum mechanics is in excellent agreement with experiment.

Except in Chapter X, we shall only consider non-relativistic quantum mechanics. Accordingly, all the following considerations, such as the Schrödinger wave equation and the equations of classical mechanics which we shall state, will only be applicable to non-relativistic phenomena. The validity of non-relativistic quantum mechanics, like that of non-relativistic classical mechanics, breaks down at energies at which the particle velocities approach the velocity of light.

5. The Schrödinger wave equation: derivation

From diffraction experiments with electrons (*cf* Born[8], particularly Chapter IV), one knows that a beam of electrons, travelling with a definite momentum of magnitude p, in the direction of the x-axis, is to be associated with a monochromatic wave of wavelength

$$\lambda = h/p \tag{1}$$

where h is Planck's constant. From dimensional considerations one sees that it has the dimensions of action, *i.e.* energy × time. Its numerical value is $6{\cdot}55 \times 10^{-27}$ erg sec. Equation (1) connects a particle concept, momentum, on the right-hand side, with a wave concept, wavelength, on the left-hand side.

If $u(x, t)$ is the amplitude of the wave at the point with coordinate x at time t [for brevity we shall simply say at (x, t)], two questions at once arise: what is the physical interpretation of $u(x, t)$, and what is the wave equation it satisfies? The 'particle' interpretation of $u(x, t)$ is due to Max Born: $|u(x, t)|^2$ *is to be interpreted as the density of electrons, i.e. the number of electrons per unit volume, at* (x, t). This postulate implies a certain normalization of the wave function $u(x, t)$. We shall deal with this below. If a monoenergetic beam of electrons of uniform intensity is to be described by a plane wave, the above postulate restricts the plane wave representation to be a complex exponential rather than a sine or cosine function:

$$u(x, t) = a \exp i\left(\frac{2\pi}{\lambda}x - \omega t\right). \tag{2}$$

(2) represents a beam of uniform density $|a|^2$; by analogy with the properties of a light quantum or photon we may expect the circular frequency ω of the wave to be related to the energy E of each electron by the equation (*cf* Born[8], particularly Chapter IV),

$$\hbar\omega = h\nu = E \tag{3}$$

where $\hbar = h/2\pi$. By considering the motion of a single electron as a wave packet we shall see that this interpretation of ω is justified.

Equation (3) again is a relation between a wave concept (ω or ν) and a particle concept (E). With the aid of equations (1) and (3) we can write the wave function (2) in the form

$$u(x,t)=a\exp\frac{i}{\hbar}(px-Et). \tag{4}$$

From equation (4) we can now derive a wave equation. By differentiation, equation (4) gives at once

$$\left.\begin{aligned}-i\hbar\frac{\partial u}{\partial x}&=pu\\ i\hbar\frac{\partial u}{\partial t}&=Eu\end{aligned}\right\} \tag{5}$$

and hence also

$$-\hbar^2\frac{\partial^2 u}{\partial x^2}=p^2u. \tag{6}$$

Using equations (5) and (6) the equation expressing conservation of energy of the electrons,

$$E=\frac{1}{2m}p^2, \tag{7}$$

can be expressed as a wave equation

$$i\hbar\frac{\partial u}{\partial t}=-\frac{\hbar^2}{2m}\frac{\partial^2 u}{\partial x^2}. \tag{8}$$

We can at once generalize the above to the case of a plane wave travelling in any direction, *i.e.* the momentum vector **p** of the electrons now has components (p_x, p_y, p_z). Equation (4) then becomes replaced by

$$u(\mathbf{r},t)=a\exp\frac{i}{\hbar}(\mathbf{pr}-Et) \tag{4a}$$

and analogously to the above,

$$\left.\begin{aligned}-i\hbar\,\mathrm{grad}\,u&=\mathbf{p}u\\ i\hbar\frac{\partial u}{\partial t}&=Eu\end{aligned}\right\} \tag{5a}$$

and

$$-\hbar^2\nabla^2u=\mathbf{p}^2u. \tag{6a}^*$$

The conservation of energy, equation (7), now gives

$$i\hbar\frac{\partial u}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2u. \tag{8a}$$

So far we have only considered free particles, not subject to any field of force, which move with uniform velocity in a fixed direction. What

---

[*] The definitions, notation and theorems of vector analysis, required in this book, are given in Appendix A.

is the correct generalization of equation (8a) if the particles move in a field characterized by a potential $V(\mathbf{r})$? (In the case of electrons, $V(\mathbf{r})$ could be an electrostatic potential.) In this case one will not expect the wave function associated with the particles to be a plane wave; for their trajectories are now curved and not straight lines and the density of particles must be dependent on the spatial coordinates. Equations (5a) which were deduced from the plane wave solution, equation (4a), need not necessarily hold in this more general case. Nevertheless, *we shall assume that equations (5a) are generally valid, even in the presence of an external field.* The classical conservation of energy equation

$$E = \frac{1}{2m}\mathbf{p}^2 + V(\mathbf{r}) \tag{7b}$$

then at once gives the general wave equation

$$i\hbar\frac{\partial u}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 u + V(\mathbf{r})u. \tag{8b}$$

Equation (8b) is the *time-dependent Schrödinger wave equation* which is fundamental to quantum mechanics.

6. The Schrödinger wave equation: discussion

We must next discuss the implications of the wave equation (5.8b).

The wave equation (5.8b) is a partial differential equation involving the first time derivative and the second-order spatial derivatives. In general, it is to be solved within some fundamental domain on whose boundaries some boundary conditions are imposed. These specify whether the system is isolated or how it interacts with the external world. Given the wave function $u(\mathbf{r}, t)$ at some initial time $t = t_0$, one can calculate it for all later times $t > t_0$ from the wave equation. Thus the Schrödinger wave equation plays the role of the equations of motion in quantum mechanics; it gives the temporal development of the system. For it will become clear that an exact knowledge of the wave function is the maximum knowledge of the system which is possible according to quantum theory: all possible information about the system can be calculated from this wave function.

We next consider equations (5.5a). In the case of a plane wave, these are quite ordinary equations in which E, the energy, and $\mathbf{p}$, the momentum vector, are ordinary numbers; (to be more accurate, the momentum vector $\mathbf{p}$ specifies a set of three numbers p_x, p_y, p_z). But as was stressed in the last section, it is one of our fundamental assumptions that equations (5.5a) are *generally valid* and not only for plane waves. The question then arises how these equations are to be interpreted in the general case.

Quantities such as position or momentum coordinates or energy, which one measures experimentally, are called *observables*. In classical physics, observables are represented by ordinary variables. In quantum mechanics, observables are represented by *operators*, *i.e.* by *quantities which operate on the wave function giving a new wave function.* Logically one must distinguish between these three concepts of 'the observable A', 'the variable A' and 'the operator A', and one should denote them by different symbols. However, it causes no confusion and it is in practice most convenient to use the same symbol for all three concepts, and we shall do so in general. Only if we especially want to stress that we refer to the operator corresponding to an observable A, shall we denote it by A_{op}.

In section 5 we had examples of such quantum mechanical operators. From equation (5.5a), for example, we can say that the observable momentum is represented by the operator $\mathbf{p}_{op}$ defined by

$$\mathbf{p}_{op} \equiv -i\hbar \text{ grad.} \tag{1}$$

The momentum operator involves differentiation and is hence called a differential operator. Like momentum itself, it is of course a vector; we shall call it a vector operator; equation (1) really stands for three equations

$$(p_x)_{op} \equiv -i\hbar \frac{\partial}{\partial x}, \textit{ etc.} \tag{1a}$$

Observables are not necessarily represented by differential operators. Thus the operator representing the position vector is simply represented by the position vector itself

$$\mathbf{r}_{op} = \mathbf{r}. \tag{2}$$

We see this, for example, from equations (5.7b) and (5.8b), where we did not replace the variable $\mathbf{r}$ in going from the classical to the quantum equation. Equation (2) implies that, generally, given a wave function $u(\mathbf{r}, t)$, the position operator defines a set of three new wave functions

$$\mathbf{r}_{op}u(\mathbf{r}, t) = \mathbf{r}u(\mathbf{r}, t)$$

again, this equation stands for three equations

$$x_{op}u(\mathbf{r}, t) = xu(\mathbf{r}, t), \textit{ etc}).$$

Finally, the energy operator E_{op} is, from equation (5.5a),

$$E_{op} \equiv i\hbar \frac{\partial}{\partial t}. \tag{3}$$

Though we gave plausibility arguments for our choice of operators the correspondences between observables and operators, as in equations (1)–(3), are of course in the nature of postulates. That these postulates lead to quantitative agreement with experimental facts is their ultimate justification. In the above we have used the suffix *op* if we wanted to stress that a quantity is regarded as an operator. But in practice no confusion or ambiguity is likely to arise, and, in general, the suffix *op* will be omitted.

6. THE SCHRÖDINGER WAVE EQUATION: DISCUSSION

Looking at the derivation of the wave equation (5.8b), we see that we started from the classical conservation of energy equation,

$$E - \frac{1}{2m}\mathbf{p}^2 - V(\mathbf{r}) = 0, \tag{5.7b}$$

replaced the classical observables by the corresponding quantum mechanical operators, equations (1) to (3), and let the resulting operator operate on the wave function. The procedure which we have here stated for obtaining the quantum-mechanical analogue of an equation of classical physics is generally valid. It enables one to write down at once quantum-mechanical equations in all problems in which a classical analogue exists. Let us, as an example, write down the wave equation of a hydrogen atom. If $\mathbf{r}_e$, $\mathbf{r}_p$, $\mathbf{p}_e$, $\mathbf{p}_p$, m_e, m_p are the position vectors, momentum vectors and masses of electron and proton respectively, the classical conservation of energy is expressed by

$$E - H(\mathbf{r}_e, \mathbf{r}_p, \mathbf{p}_e, \mathbf{p}_p) \equiv E - \frac{1}{2m_e}\mathbf{p}_e^2 - \frac{1}{2m_p}\mathbf{p}_p^2 + \frac{e^2}{r} = 0, \tag{4}$$

where $r = |\mathbf{r}_e - \mathbf{r}_p|$ is the distance between proton and electron and $-e$ is the electronic charge. The function H, introduced in (4), is called the *Hamiltonian* function of the system, *i.e.* the energy expressed in terms of the position and momentum coordinates. From (4) we obtain at once the wave equation of the hydrogen atom. It is now an equation for a wave function $u = u(\mathbf{r}_p, \mathbf{r}_e, t)$:

$$i\hbar\frac{\partial u}{\partial t} = -\frac{\hbar^2}{2m_p}\nabla_p^2 u - \frac{\hbar^2}{2m_e}\nabla_e^2 u - \frac{e^2}{r}u,$$

where the Laplacian operators ∇_p^2 and ∇_e^2 operate on the proton and electron coordinates of $u(\mathbf{r}_p, \mathbf{r}_e, t)$ respectively, *i.e.*

$$\nabla_p^2 \equiv \frac{\partial^2}{\partial x_p^2} + \frac{\partial^2}{\partial y_p^2} + \frac{\partial^2}{\partial z_p^2}, \textit{ etc.}$$

We shall only in Chapter IV deal with the interpretation of the above formalism. However, we shall anticipate that discussion in order to be able to introduce some of the physical interpretation already at this stage. It is one of the interpretative axioms of quantum mechanics, that if a system is described by a wave function u, and the effect of operating on u with an operator A_{op}, corresponding to an observable A, is equivalent to multiplying the wave function by a number a, *i.e.*

$$A_{op}u = au, \tag{5}$$

then, in the state of the system described by the wave function u, the observable A has the value a. Equation (5) defines an eigenvalue problem of the kind discussed in Chapter I. We call u an eigenfunction and a the corresponding eigenvalue of the operator A_{op}. For the special case of the momentum of the electrons of an electron beam, equation (5) would read

$$\mathbf{p}_{op}u = \mathbf{p}u. \tag{5'}$$

We see that the plane wave, equation (5.4a),

$$u = a \exp \frac{i}{\hbar}(\mathbf{pr} - Et), \tag{5.4a}$$

satisfies (5′), *i.e.* it represents a beam of electrons with momentum $\mathbf{p}$, as assumed earlier.

We next want to investigate the interpretation of $|u(x, t)|^2$ as a particle density. Let us, to begin with, consider a fixed time, $t = t_0$, so that we can forget about the time dependence and simply write $u = u(\mathbf{r})$. If our system consists of N particles, we require that

$$\int |u(\mathbf{r})|^2 d\tau = N \tag{6}$$

fixed time

the integration being over the whole volume occupied by the system; this may be the whole of space, or if the particles are confined to a 'box' it may be the volume of this box. In the case of a plane wave, equation (5.4a), such a normalization of the wave function is not possible. A plane wave is infinitely long and has an infinitely wide wave front. The integral in (6) diverges, since it is to be taken over all space, and the integrand is constant. Nevertheless a plane wave represents a useful mathematical idealization, just like strictly monochromatic waves leading to infinitely narrow spectral lines in optics (*cf* Exercise I.6). For a plane wave representation to make sense, we simply require the wave front and the length of the wave train to be large compared with the wavelength λ. But if we use such a plane wave idealization we must alter the normalization (6). A physically reasonable normalization is to norm the wave function to give the correct density ϱ of the beam of particles

$$\int_{\substack{\text{unit}\\\text{volume}}} |u(\mathbf{r})|^2 d\tau = \int_{\substack{\text{unit}\\\text{volume}}} \left|a \exp \frac{i}{\hbar}(\mathbf{pr} - Et)\right|^2 d\tau = |a|^2 = \varrho. \tag{7}$$

This determines a apart from a factor of modulus unity. Now we can, without loss of generality, take a to be real. For we shall presently see that multiplying a wave function by a factor of modulus unity (*e.g.* $e^{i\theta}$, where θ is any real number) does not affect the physical predictions which one makes from the wave function.

We have interpreted $|u(\mathbf{r}, t)|^2$ as the particle density at $(\mathbf{r}, t)$. This interpretation must certainly fail if $N = 1$. $|u(\mathbf{r}, t)|^2 d\tau$ is now to be interpreted as the probability of finding the particle in the volume element $d\tau$ at the point $\mathbf{r}$ at time t. This interpretation of $|u(\mathbf{r}, t)|^2$, due to Born, is much more radical than the original interpretation as a particle density. It implies that the position of a single particle at a definite time t is not uniquely determined, but that there exists only a probability distribution of finding it at any particular point $\mathbf{r}$. The laws of quantum mechanics are of a statistical nature, even if we deal with only a single particle.

For a single particle we replace (6) by

$$\int |u(\mathbf{r})|^2 d\tau = 1. \tag{6a}$$

One cannot, in general, describe a single particle by a plane wave. A plane wave implies equal probability of finding the particle anywhere, whereas in practice a single particle is to some degree localized, *i.e.* one has some knowledge of its position. It is more reasonable to describe it by a wave packet*. A wave packet can be thought of as a superposition of a group of monochromatic waves whose momenta lie in a small interval

$$\mathbf{p}_0 < \mathbf{p} < \mathbf{p}_0 + \Delta\mathbf{p}_0.$$

At some point P, the individual sinusoidal waves are all in phase and build up to a large resultant amplitude. Outside a region

$$\mathbf{r}_0 < \mathbf{r} < \mathbf{r}_0 + \Delta\mathbf{r}_0,$$

surrounding the point P, the individual waves interfere destructively. From the Fourier integral theorem one knows that increasing the size of the momentum interval $\Delta\mathbf{p}_0$ of the group of plane waves enables one to decrease the extent $\Delta\mathbf{r}_0$ of the wave packet. A plane wave represents the extreme case of a particle whose position is completely unknown but whose momentum is defined with absolute accuracy. Superposing a large number of plane waves (*i.e.* large $\Delta\mathbf{p}_0$), so that the momentum is now badly defined, enables one to restrict $\Delta\mathbf{r}_0$ to a narrow interval, *i.e.* the position of the particle is well defined. We have here a first example of Heisenberg's *uncertainty principle*, which states that it is not possible to have an accurate knowledge of the simultaneous position and momentum of a particle.

For a monochromatic wave

$$u = e^{i\varphi}, \quad \varphi = \frac{2\pi}{\lambda}x - \omega t,$$

we define a *phase velocity* c_φ by

$$c_\varphi = \left(\frac{\partial x}{\partial t}\right)_{\varphi=\text{const.}} = \frac{\omega\lambda}{2\pi} = \nu\lambda;$$

c_φ is the velocity with which a particular point of the wave, say one of the crests, moves along. In general the phase velocity c_φ depends on the wavelength

---

* In practice, one usually deals with beams of many particles. For these, one can usually approximate the actual wave function by a plane wave, as explained above, so that one can avoid the somewhat complicated analysis in terms of wave packets.

$$c_\varphi = c_\varphi(\lambda);$$

this is the well-known phenomenon of dispersion. As a result of this dispersion a wave packet does not retain its shape accurately, but if it is a narrow wave packet it does retain its form approximately. We can then talk of a *group velocity* c_g of the wave packet, which is the velocity with which this pulse propagates itself. It is not difficult to show that in the limiting case of a sufficiently narrow wave packet the group velocity is given by

$$c_g = \frac{d\nu}{d(1/\lambda)} \tag{8}$$

(*cf* Exercise **II**.3). If we identify the group velocity c_g with the particle velocity v we have, using equations (8) and (5.1),

$$v = \frac{d\nu}{d(1/\lambda)} = \frac{h\,d\nu}{m\,dv}$$

and integrating this equation,

$$\hbar\omega = h\nu = \tfrac{1}{2}mv^2 = E. \tag{5.3}$$

We have thus obtained an independent verification that equation (5.3) gives a consistent interpretation of ω in terms of the energy.

So far we have considered one fixed instant of time $t = t_0$. We must now turn to the time dependence of the wave function. If we exclude processes in which particles are annihilated or created, the normalization integral, equation (6), must remain constant throughout time,

$$\frac{\partial N}{\partial t} = 0. \tag{9}$$

Now

$$\frac{\partial}{\partial t}|u|^2 = \bar{u}\frac{\partial u}{\partial t} + \frac{\partial \bar{u}}{\partial t}u. \tag{10}$$

From the Schrödinger equation

$$i\hbar\frac{\partial u}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 u + V(\mathbf{r})u$$

and its complex conjugate

$$-i\hbar\frac{\partial \bar{u}}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2 \bar{u} + V(\mathbf{r})\bar{u}$$

(we are always assuming that the potential is real), we obtain, multiplying by $\bar{u}$ and u respectively and subtracting,

$$\frac{\partial}{\partial t}|u|^2 = -\frac{\hbar}{2mi}(\bar{u}\nabla^2 u - u\nabla^2 \bar{u}) \tag{11}$$

If we define a vector $\mathbf{j}$ by the equation

$$\mathbf{j}=\frac{\hbar}{2mi}(\bar{u}\text{ grad }u-u\text{ grad }\bar{u}) \tag{12}$$

and integrate (11) over a volume V bounded by a surface S, we obtain, using Green's theorem (Appendix A, equation (A11)),

$$\frac{\partial}{\partial t}\int_V |u|^2 d\tau=-\int_V \operatorname{div}\mathbf{j}\,d\tau=-\int_S j_n dS \tag{13}$$

where $d\tau$ is a volume element within the volume V, dS a surface element on the surface S, and j_n the component of the vector $\mathbf{j}$ in the direction of the outward normal $\mathbf{n}$ of the surface element dS. We interpret $\mathbf{j}$ as a current vector, *i.e.* $j_n(\mathbf{r}, t)dS$ is the rate at which particles cross a surface element dS (normal to the vector $\mathbf{n}$) situated at $\mathbf{r}$, at time t. Equation (13) expresses the fact that the rate of decrease of particles within the volume V equals the rate at which they escape across the boundary surface S. If $\mathcal{N}=1$, $\mathbf{j}$ is the probability current vector; $j_n dS$ defines the probability of an electron crossing an element of area dS in the direction $\mathbf{n}$ normal to dS per unit time.

If our system is localized, we can choose the surface S so that $u=0$ everywhere on and outside S, *i.e.* from (12) $\mathbf{j}=0$. Then the right-hand side of (13) vanishes and equation (9) follows. Even if $u(\mathbf{r}, t)$ extends throughout space, equation (9) remains true, provided $u(\mathbf{r}, t)$ has a finite normalization integral. For, in this case, $u \to 0$ sufficiently rapidly at infinity for the right-hand side of (13) again to vanish*. Thus the normalization persists whenever the normalization integral $\mathcal{N}$ is finite. But these are the only cases of interest; for if $\mathcal{N}$ is infinite, it is meaningless to ask for its rate of change, nor can such wave functions represent an actual physical system.

7. Energy eigenfunctions: theory

For a system in a state with definite energy E, we can put

$$u(\mathbf{r}, t)=\psi(\mathbf{r})e^{-iEt/\hbar}. \tag{1}$$

For it follows from the discussion of equation (6.5) and from equation (6.3) defining the energy operator that a measurement of the energy of the system in the state (1) necessarily gives the result E. We have

$$i\hbar\frac{\partial u}{\partial t}=Eu$$

and the Schrödinger equation (5.8b) becomes

$$H\psi\equiv\left(-\frac{\hbar^2}{2m}\nabla^2+V\right)\psi=E\psi. \tag{2}$$

---

* These conditions are mathematically not quite general enough, but they are sufficient for the functions one meets in quantum mechanics.

Equation (2) is called the time-independent Schrödinger equation. It represents a typical eigenvalue problem of quantum mechanics. The solutions represent stationary energy states, the energy being one of the possible eigenvalues of (2); the corresponding wave functions are called energy eigenfunctions. To define the eigenvalue problem completely, we must specify the fundamental domain V within which (2) is to be solved and the boundary conditions which ψ must satisfy on the boundary S of V. In Chapter III we shall solve some energy eigenvalue problems and we shall then discuss these boundary conditions in more detail. For the present we want to discuss some general aspects of (2), assuming always that suitable boundary conditions are imposed.

(i) *Discreteness*

The eigenvalue problem defined above is very general, and it does not necessarily possess a discrete spectrum of eigenvalues. We want to restrict ourselves, for the time being, to a discrete spectrum of eigenvalues*

$$E_1 < E_2 < \ldots < E_n < \ldots \tag{3}$$

and corresponding eigenfunctions

$$\psi_1, \psi_2, \ldots \psi_n, \ldots \tag{4}$$

which we take as normalized:

$$(\psi_n, \psi_n) = 1. \tag{5}$$

That the ψ_n can be normed depends on a general theorem of eigenvalue theory: *if an eigenvalue problem has a completely discrete spectrum the eigenfunctions can be normed, and vice versa.* This theorem does not help us to decide whether a particular eigenvalue problem has a discrete spectrum or not. However, in the following important classes of cases the spectrum is always discrete:

(*a*) if the system is localized, *i.e.* there is a region V bounded by a surface S such that $\psi = 0$ on and outside S;

(*b*) for problems with infinite fundamental domains, provided the potential $V(\mathbf{r}) \to \infty$ whenever $|\mathbf{r}| \to \infty$†. We can think of this case as a quasi-localized problem; the potential increasing at large distances prevents any particles from escaping to large distances.

From the general theorem stated above, it follows that for problems with infinite fundamental domains and discrete spectra, $\psi(\mathbf{r}) \to 0$ sufficiently rapidly at infinity for $\psi(\mathbf{r})$ to have a finite norm. If $\psi(\mathbf{r})$ and $\chi(\mathbf{r})$ are two functions with finite norms, V is a finite domain bounded by a closed surface S, and $\mathbf{n}$ is the normal to an element of surface dS, then the surface integral

---

* One can show that there always exists a lowest energy eigenvalue, corresponding to the ground state of the system.

† For criteria (*a*) and (*b*) to apply, the potential $V(\mathbf{r})$ must not have too strong negative singularities. We shall not state these conditions explicitly, as they are usually satisfied in quantum mechanics.

$$\int_S \psi(\mathbf{r})\frac{\partial}{\partial n}\chi(\mathbf{r})dS$$

vanishes as we let the surface S move off to infinity in *all* directions, (*Cf* footnote on page 31.)

(*ii*) *Reality*

The eigenvalues of the above eigenvalue problem are all real. If E is an eigenvalue and ψ the corresponding eigenfunction, we have from (2) and (5) that

$$E = -\frac{\hbar^2}{2m}\int_V \bar{\psi}\nabla^2\psi d\tau + \int_V V|\psi|^2 d\tau. \tag{6}$$

In (6), the second integral is real, and we can transform the first by means of equation (A 10) of Appendix A, to give

$$-\int_V \text{grad}\ \bar{\psi}\ \text{grad}\ \psi d\tau + \int_S \bar{\psi}\frac{\partial\psi}{\partial n}dS.$$

The first of these integrals is real and the second vanishes, as we saw in the last paragraph. Hence, the eigenvalues are real.

(*iii*) *Orthogonality*

If E_m and E_n are two different eigenvalues of (2), then any eigenfunctions ψ_m and ψ_n belonging to these eigenvalues respectively are orthogonal, *i.e.*

$$(\psi_m, \psi_n) = 0, \quad m \neq n. \tag{7}$$

From the Schrödinger equation (2) we have

$$\bar{\psi}_m\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi_n = E_n\bar{\psi}_m\psi_n;$$

interchanging m and n, taking complex conjugates and using the reality of the eigenvalues, we obtain

$$\psi_n\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\bar{\psi}_m = E_m\bar{\psi}_m\psi_n.$$

Subtracting these two equations, integrating over the fundamental domain V and using equation (A 11) of Appendix A, we obtain

$$
\begin{aligned}
(E_n - E_m)(\psi_m, \psi_n) &= -\frac{\hbar^2}{2m}\int_V (\bar{\psi}_m \nabla^2 \psi_n - \psi_n \nabla^2 \bar{\psi}_m)\, d\tau \\
&= -\frac{\hbar^2}{2m}\int_S \left(\bar{\psi}_m \frac{\partial \psi_n}{\partial n} - \psi_n \frac{\partial \bar{\psi}_m}{\partial n}\right) dS \\
&= 0. \qquad (8)
\end{aligned}
$$

Since $E_n \neq E_m$, this equation implies $(\psi_m, \psi_n) = 0$.

(iv) Completeness

Suppose we know the wave function $u(\mathbf{r}, t_0)$ at some time t_0. The time-dependent Schrödinger equation then completely determines the wave function $u(\mathbf{r}, t)$ at later times $t > t_0$. If $u(\mathbf{r}, t_0)$ can be written in the form

$$u(\mathbf{r}, t_0) = \sum_{n=1}^{\infty} a_n \psi_n(\mathbf{r}) \qquad (9)$$

then we have at once the general solution

$$u(\mathbf{r}, t) = \sum_{n=1}^{\infty} a_n \psi_n(\mathbf{r}) e^{-iE_n(t-t_0)/\hbar}, \qquad (10)$$

as this function satisfies the wave equation (2) and the initial conditions (9). Given that $u(\mathbf{r}, t_0)$ is a linear superposition of energy eigenfunctions, equation (9) completely determines the coefficients a_n:

$$a_n = (\psi_n(\mathbf{r}), u(\mathbf{r}, t_0)) = \int_V \overline{\psi_n(\mathbf{r})}\, u(\mathbf{r}, t_0)\, d\tau. \qquad (11)$$

The importance of the above analysis is due to the fact that the energy eigenfunctions form a complete set: any arbitrary function $f(\mathbf{r})$ with finite norm can be expanded in an infinite series of the form (9). Hence equation (10) gives a general solution of the time-dependent problem for all physically admissible initial conditions.

EXERCISES II

II.1. Derive the angular momentum operator $\mathbf{L}_{op}$ of a system of n point masses whose (classical) position and momentum coordinates are $\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_n$ and $\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_n$ respectively.

II.2. Write down the Schrödinger equation for the helium atom.

II.3. If one represents a one-dimensional wave packet by a Fourier integral

$$u(x, t) = \int_{p_0 - \Delta p_0/2}^{p_0 + \Delta p_0/2} a(p) \exp\left[\frac{i}{\hbar}(px - Et)\right] dp \qquad (\text{II.3})$$

where a and E are continuous functions of p, derive equation (6.8) for the group velocity of an infinitely sharp wave packet (*i.e.* $\Delta p_0 \to 0$).

II.4. If $\varrho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ are the particle density and particle current vector at $(\mathbf{r}, t)$ respectively, prove the 'continuity equation'

$$\frac{\partial \varrho}{\partial t} + \operatorname{div} \mathbf{j} = 0. \qquad \text{(II. 4)}$$

II.5. Suppose equation (I.6a) (Exercise I.6), describes a single particle, *i.e.*

$$\int_{-\infty}^{\infty} |f(x)|^2 dx = 1.$$

Verify Heisenberg's uncertainty principle for this wave function; *i.e.* if Δx and Δp are the uncertainties in the knowledge of the position and momentum of the particle, then $(\Delta x)(\Delta p)$ is of the order of magnitude of Planck's constant h:

$$(\Delta x)(\Delta p) \simeq h.$$

(*N.B.* In Chapter IV we shall consider the uncertainty principle more critically.)

III

ENERGY EIGENFUNCTIONS: EXAMPLES

ONLY a few of the energy eigenvalue problems which occur in quantum mechanics admit simple rigorous solutions. Yet these problems are of great importance, as they can in many cases be taken as the basis of an approximate method of solution.

We shall, however, not deal exhaustively with these 'standard problems' which can, for example, be found in the books by Schiff[12] or Sommerfeld[13]. We only want to deal with a few of them to make the reader more familiar with the general theory so far developed. The examples selected will be fundamental to our treatment of general topics, such as angular momentum, scattering theory, *etc*, in later chapters.

8. One-dimensional potential barrier

For a one-dimensional problem the time-independent Schrödinger equation (7.2) takes the form

$$\frac{d^2\psi(x)}{dx^2}+\frac{2m}{\hbar^2}[E-V(x)]\psi(x)=0. \tag{1}$$

Suppose $V(x)$ represents a potential barrier of finite height

$$\left.\begin{aligned} V(x)&=0, & x&<0 \\ V(x)&=V_0>0, & x&>0 \end{aligned}\right\} \tag{2}$$

as shown in *Figure 3*, and that we have a beam of particles incident from the left-hand side, $x<0$, encountering the potential barrier. This eigenvalue problem has a completely continuous spectrum of energy eigenvalues $E>0$.

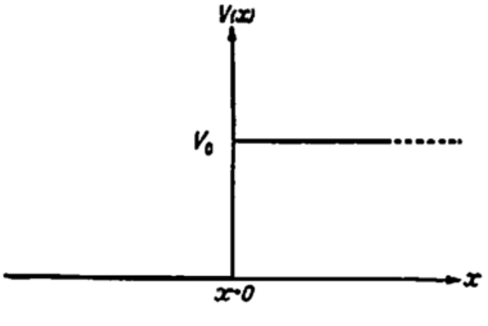

Figure 3. One-dimensional potential barrier

At $x=0$, $V(x)$ is discontinuous. But since E, $V(x)$ and $\psi(x)$ are everywhere finite ($\psi(x)$ so that the probability interpretation of $|\psi|^2$

should have a meaning), it follows from (1) that $\psi''(x)$ is everywhere finite. Hence $\psi'(x)$ and, *a fortiori*, $\psi(x)$ are everywhere finite and continuous, including $x=0$. These continuity conditions are generally valid. Given a general (three-dimensional) surface S on which the potential $V(\mathbf{r})$ experiences a finite jump discontinuity, the wave function ψ and its normal derivative $\partial\psi/\partial n$ ($\partial\psi/\partial n$ denotes differentiation along the normal to S) must be continuous at every point of S. (The tangential derivatives are in any case continuous so that we can say ψ and gradψ are everywhere continuous.) Hence the density $|\psi|^2$ and the current $\mathbf{j}$, equation (6.12), are everywhere finite and continuous, a physically reasonable condition.

The solution of (1) and (2) differs considerably according as $E>V_0$ or $E<V_0$.

If $E>V_0$, we write

$$k_1 = +\left(\frac{2m}{\hbar^2}E\right)^{\frac{1}{2}}, \quad k_2 = +\left(\frac{2m}{\hbar^2}(E-V_0)\right)^{\frac{1}{2}}, \tag{3}$$

and obtain the solution

$$\psi(x) = \psi_1(x) = A\mathrm{e}^{ik_1x} + B\mathrm{e}^{-ik_1x} \quad (x<0) \tag{4a}$$

$$\psi(x) = \psi_2(x) = C\mathrm{e}^{ik_2x} \qquad (x>0). \tag{4b}$$

No term e^{-ik_2x} occurs in ψ_2 as no waves incident from $x=+\infty$ are physically allowed in our problem. From the continuity conditions

$$\psi_1(0) = \psi_2(0), \quad \psi_1'(0) = \psi_2'(0) \tag{5}$$

we obtain

$$\frac{B}{A} = \frac{k_1-k_2}{k_1+k_2}, \quad \frac{C}{A} = \frac{2k_1}{k_1+k_2}. \tag{6}$$

Now e^{ik_1x} and e^{-ik_1x} correspond to an incident and a reflected wave respectively, the incident and reflected currents being, according to (6.12),

$$j_I = \frac{\hbar k_1}{m}|A|^2, \quad j_R = \frac{\hbar k_1}{m}|B|^2; \tag{7}$$

equation (4b) defines a transmitted beam, the transmitted current being

$$j_T = \frac{\hbar k_2}{m}|C|^2. \tag{8}$$

We define the reflection and transmission coefficients

$$R = \frac{j_R}{j_I} = \left|\frac{B}{A}\right|^2 = \left(\frac{k_1-k_2}{k_1+k_2}\right)^2 \tag{9}$$

$$T = \frac{j_T}{j_I} = \frac{k_2}{k_1}\left|\frac{C}{A}\right| = \frac{4k_1k_2}{(k_1+k_2)^2}, \tag{10}$$

and $R+T=1$, as expected.

If $E<V_0$, the solution of the eigenvalue problem is still given by equations (4) to (6). But now $k_2=i\kappa$ is purely imaginary. In the region $x>0$, the solution $e^{-ik_2x}=e^{\kappa x}$ is still excluded as it diverges as $x \to \infty$. Equation (7) still holds, but, from (6.12), (8) is replaced by $j_T=0$, and the reflection and transmission coefficients become

$$R=\frac{j_R}{j_I}=\left|\frac{B}{A}\right|^2=\left|\frac{k_1-i\kappa}{k_1+i\kappa}\right|^2=1 \tag{11}$$

$$T=\frac{j_T}{j_I}=0. \tag{12}$$

Thus all the particles are reflected, as one would also expect classically in the case $E<V_0$. On the other hand, the result for the case $E>V_0$ differs from that of classical mechanics. Classically one requires $T=1$, $R=0$, if $E>V_0$.

If we consider the limit as $V_0 \to \infty$, keeping E finite, then $\kappa \to \infty$, but k_1 remains finite, and in the limit

$$\frac{B}{A}=-1, \quad \frac{C}{A}=0,$$

i.e.

$$\begin{aligned} \psi&=\psi_1=a\sin k_1x \quad &(x<0)\\ \psi&=\psi_2=0 &(x>0) \end{aligned}$$

a being an arbitrary constant. Thus, the boundary conditions at a boundary where there is an infinite potential barrier are that the wave function is zero (and is continuous), whereas the derivative ψ' is undetermined (and is, in general, discontinuous). Physically, $V_0=\infty$ means that the barrier is completely impenetrable, the particles are always confined to the region $x<0$. This result can clearly be generalized.

9. One-dimensional square well potential with rigid walls

We consider the one-dimensional motion of a particle restricted to the region $|x|<L$, *i.e.* the particle is inside a one-dimensional box or square well bounded by impenetrable perfectly rigid walls at $x=\pm L$. According to section 8, the corresponding potential $V(x)$ is

$$\begin{aligned} V(x)&=0, \quad &|x|<L\\ V(x)&=\infty &|x|>L \end{aligned}$$

and the corresponding eigenvalue problem is the same as the eigenvalue problem

$$\psi''(x)+k^2\psi(x)=0 \tag{1}$$

$$k=+\left(\frac{2m}{\hbar^2}E\right)^{\frac{1}{2}} \tag{2}$$

for the region $-L<x<L$, with the boundary conditions

$$\psi(-L)=\psi(L)=0. \tag{3}$$

From section 7(*i*) we know that this eigenvalue problem has a completely discrete spectrum of eigenvalues. The general solution of (1) is

$$\psi = A\cos kx + B\sin kx \tag{4}$$

and the boundary conditions (3) give at once

$$A\cos kL = B\sin kL = 0. \tag{5}$$

Hence we have two possible types of solution, either

$$(i)\ A=0,\quad k_n=\frac{n\pi}{2L},\quad \psi_n=B_n\sin\frac{n\pi}{2L}x,\quad n=2,4,6,\ldots \tag{6a}$$

or

$$(ii)\ B=0,\quad k_n=\frac{n\pi}{2L},\quad \psi_n=A_n\cos\frac{n\pi}{2L}x,\quad n=1,3,5,\ldots \tag{6b}$$

the coefficients A_n and B_n being determined from the normalization of the wave functions. The corresponding energy eigenvalues are given by the same expression for both types of solution,

$$E_n=\frac{\hbar^2}{2m}k_n{}^2=\frac{1}{2m}\left(\frac{\hbar\pi}{2L}\right)^2 n^2,\quad (n=1,2,3,\ldots). \tag{7}$$

Thus the energy eigenvalues form an increasing infinite sequence of discrete energy levels.

In *Figure 4* the first few eigenfunctions are shown (with arbitrary normalizations). *Figure 4* illustrates the fact, easily verified from equations (6), that the nth eigenfunction (ordered so that $E_1<E_2<\ldots$) has exactly $(n-1)$ zeros (or nodes) inside the fundamental domain. This theorem is generally valid for one-dimensional eigenvalue problems with discrete spectra. But no analogous theorem about nodal lines or surfaces exists for partial differential equations. The above eigensolutions are non-degenerate. This again is generally true for eigenvalue problems comprising an ordinary second-order differential equation and two boundary conditions. For the two boundary conditions uniquely determine the two arbitrary constants which occur in the general solution of the differential equation, *i.e.* the eigensolution.

The two types of solution, equations (6), show an important difference. The sine and cosine solutions have different *parity*. Generally, we define the parity of a function $f(x_1, y_1, z_1, x_2, \ldots, x_n, y_n, z_n)$ of $3n$ spatial coordinates $x_1, y_1, z_1, x_2, \ldots x_n, y_n, z_n$, as follows. We consider an *inversion* of the coordinate system, *i.e.* a reflection of the coordinate axes in the origin. A point with position vector $\mathbf{r}$ has the position vector $-\mathbf{r}$ after inversion. Thus, under inversion, the function $f(x_1, y_1, \ldots z_n)$ becomes transformed into $f(-x_1, -y_1, \ldots -z_n)$. We say that $f(x_1, \ldots z_n)$ is an *even* function and has *parity* $+1$, if

$$f(-x_1, \ldots -z_n)=f(x_1, \ldots z_n), \tag{8a}$$

and it is an *odd* function and has *parity* -1, if

$$f(-x_1, \ldots -z_n)=-f(x_1, \ldots z_n). \tag{8b}$$

We see that the sine and cosine solutions are odd and even, and have parities -1 and $+1$ respectively.

In general, a function will be neither odd nor even, but a linear combination of the two types of function. But the interactions with which one deals in quantum mechanics always have a definite parity (in our above example, $V(-x)=V(x)$, hence $V(x)$ is an even function and has parity $+1$), and the corresponding wave functions then also always have a definite parity [*cf* Exercise III.1]. Different states can then be classified according to their parity. The importance of this classification is due to the fact that, when considering a particular atomic or nuclear process, states with different parity behave quite differently. We shall consider these questions in much greater detail later on.

10. Free particle wave functions and three-dimensional square well potential

In section 5 we obtained as the wave function of a particle of mass m in free space

$$\psi(\mathbf{r})=e^{i\mathbf{k}\mathbf{r}}, \tag{1}$$

where the momentum $\mathbf{p}$ and the wave vector $\mathbf{k}$ are related by

$$\mathbf{p}=\hbar\mathbf{k}. \tag{2}$$

The wave function $\psi(\mathbf{r})$ is an eigenfunction of momentum and energy, with eigenvalues $\mathbf{p}$ and $(\mathbf{p}^2/2m)$ respectively, for all values of the momentum vector $\mathbf{p}$, since

$$-i\hbar \operatorname{grad} \psi=\mathbf{p}\psi \tag{3a}$$

$$-\frac{\hbar^2}{2m}\nabla^2\psi=\frac{1}{2m}\mathbf{p}^2\psi. \tag{3b}$$

That $\psi(\mathbf{r})$ is a simultaneous eigenfunction of momentum and energy is due to the fact that for a free particle the energy H is a function of the momentum $\mathbf{p}$ only,

$$H=\frac{1}{2m}\mathbf{p}^2;$$

as in classical physics, a knowledge of the momentum completely determines the energy under these conditions.

The eigenfunctions (1) are mutually orthogonal, as one sees at once by using the Dirac formalism for eigenvalue problems with continuous spectra. In fact, from equations (4.13), (4.16) and (4.18) one has at once

$$(e^{i\mathbf{k}'\mathbf{r}}, e^{i\mathbf{k}\mathbf{r}})=\int_{\substack{\text{All}\\ \text{space}}} e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}}d\tau=(2\pi)^3\ \delta(\mathbf{k}-\mathbf{k}'). \tag{4}$$

It is, however, possible to avoid introducing these continuum eigenfunctions and to force the problem to have a discrete spectrum. This is achieved by enclosing the system in a sufficiently large, but finite, box and imposing suitable boundary conditions on the walls of the box. Since any actual physical system is always to some extent localized, it is clear that introducing a container the dimensions of which are sufficiently large compared to those of significance for the system considered, will affect the system to an arbitrarily small degree. Thus, we can have a 'physically plane wave' inside a box, provided the box is very large compared with the wavelength of the wave, as we saw in example I.6 and in section 6.

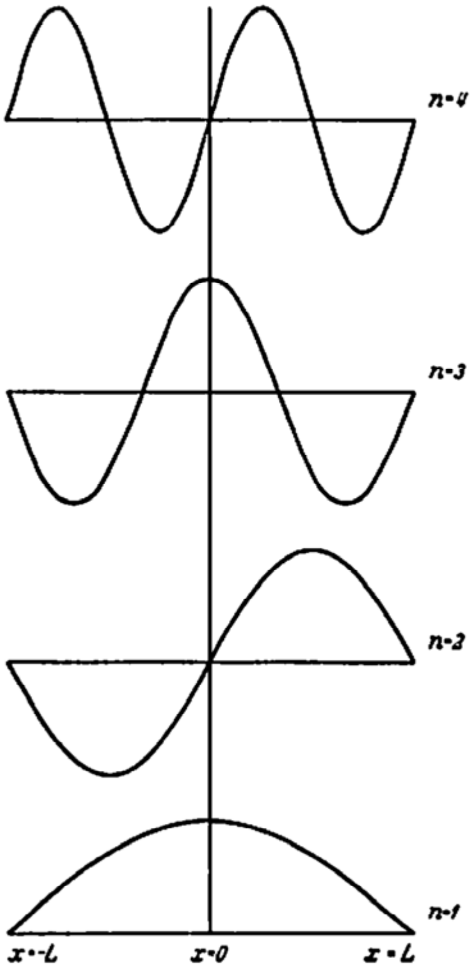

Figure 4. Energy eigenfunctions ψ_n of the one-dimensional square well potential with rigid walls, for $n=1, 2, 3, 4$

A physically natural boundary condition to assume is that the wave function vanishes on the walls of the container. A mathematically more convenient representation is obtained if one takes the box as a cube of side of length L and imposes periodic boundary conditions on the wave function. The eigenvalue problem of a free particle then becomes

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = 0, \quad \begin{pmatrix} 0<x<L \\ 0<y<L \\ 0<z<L \end{pmatrix} \tag{5}$$

where
$$k^2 = \frac{2m}{\hbar^2}E, \tag{6}$$

together with the boundary conditions that $\psi(\mathbf{r})$ be periodic with period L, *i.e.*

$$\psi(0, y, z) = \psi(L, y, z), \textit{ etc.} \tag{7}$$

The normed eigensolutions of this eigenvalue problem are easily obtained by separating the variables. They are

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{3/2}} \exp(i\mathbf{kr}), \tag{8}$$

where **k** is the vector with components (k_x, k_y, k_z) given by

$$k_x = \frac{2\pi}{L}n_1, \quad k_y = \frac{2\pi}{L}n_2, \quad k_z = \frac{2\pi}{L}n_3, \tag{9}$$

n_1, n_2 and n_3 being restricted to the natural numbers $0, \pm 1, \pm 2, \ldots$, so that the periodicity condition (7) should hold. If k is the magnitude of the vector **k**, *i.e.*

$$k = |\mathbf{k}| = \sqrt{(k_x^2 + k_y^2 + k_z^2)} = \left(\frac{2\pi}{L}\right)\sqrt{(n_1^2 + n_2^2 + n_3^2)}, \tag{10}$$

the energy eigenvalue of the state (8) is given by equation (6).

Comparing equations (1) and (8) illustrates our general discussion that introducing a sufficiently large box does not affect a physical problem. Though equation (5) has a discrete spectrum, whereas (1) has a continuous spectrum, the different momentum eigenfunctions (8) lie arbitrarily closely together if L is sufficiently large; that is, any eigenfunction of (1) can be approximated arbitrarily closely.

When dealing with a continuum of states (whether one describes these by a continuum of eigenfunctions or by a discrete set, as above) one is not in general interested in individual states, but only in groups of neighbouring states lying close together. To distinguish individual states would require measurements of absolute accuracy. Measurements of only limited accuracy enable one to say that a system is in one of a small group of neighbouring states. One then requires the number of states in such a group.

From the above analysis one can find this density of states. As in classical physics, it is often convenient to introduce *momentum space*, the rectangular coordinates of which are the components of momentum (p_x, p_y, p_z). Each point of this space corresponds to a definite momentum vector. It follows from (9) that the quantum-mechanically allowed momenta form a lattice of points in momentum space with coordinates

$$\frac{2\pi\hbar}{L}n_1, \quad \frac{2\pi\hbar}{L}n_2, \quad \frac{2\pi\hbar}{L}n_3.$$

Hence the volume per point, *i.e.* the volume per state is $(2\pi\hbar/L)^3$ and the number of states in an elementary volume $d\mathbf{p}$ of momentum space is

$$\varrho(\mathbf{p})d\mathbf{p} = \left(\frac{L}{2\pi\hbar}\right)^3 d\mathbf{p}. \tag{11}$$

Introducing polar coordinates (p, θ, φ), defined by

$$p_x = p \sin\theta\cos\varphi, \quad p_y = p\sin\theta\sin\varphi, \quad p_z = p\cos\theta, \tag{12}$$

we obtain the number of states lying in the elementary volume

$$d\mathbf{p} = p^2 dp d\Omega = p^2 dp \sin\theta \, d\theta d\varphi.$$

From (11), it is

$$\varrho(\mathbf{p})d\mathbf{p} = \varrho(\mathbf{p})p^2 dp d\Omega = \left(\frac{L}{2\pi\hbar}\right)^3 p^2 dp d\Omega = \left(\frac{L}{2\pi}\right)^3 k^2 dk d\Omega \tag{13}$$

where, as is often convenient, we have introduced the wave vector $\mathbf{k} = \mathbf{p}/\hbar$. (13) gives the number of states for which the momentum vector $\mathbf{p}$ lies within the element of solid angle $d\Omega = \sin\theta \, d\theta d\varphi$ in the direction (θ, φ) and its magnitude lies in the interval $(p, p+dp)$.

We shall subsequently see that the length L occurring in (13) cancels out identically in all physically significant quantities. From the way we introduced it this was to be expected and indeed had to be so in order that the above treatment should make sense. For this reason one often talks of the density of states per unit volume $\varrho(\mathbf{p})/L^3$ which is independent of L and, from (11), is simply $1/h^3$.

11. Linear harmonic oscillator

We consider the one-dimensional motion of a point particle of mass m in the potential field

$$V(x) = \tfrac{1}{2}\mu x^2. \tag{1}$$

According to classical mechanics, this particle experiences the force

$$-\frac{\partial V}{\partial x} = -\mu x$$

and performs simple harmonic motion with circular frequency ω_0 given by

$$\omega_0 = \left(\frac{\mu}{m}\right)^{\frac{1}{2}}. \tag{2}$$

The corresponding quantum mechanical problem consists of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\psi''(x) + \tfrac{1}{2}\mu x^2\psi(x) = E\psi(x) \tag{3}$$

and the boundary conditions,

$$\psi(x) \to 0, \quad \text{as } x \to \pm\infty. \tag{4}$$

As stated in section 7(*i*), this problem has a completely discrete spectrum of eigenvalues, since $V(x) \to \infty$ as $x \to \pm\infty$.

To solve (3), we first eliminate the constants $\hbar^2/2m$ and μ by introducing new units. We let

$$\left.\begin{aligned} \alpha^4 = \frac{m\mu}{\hbar^2}, \qquad \lambda = \frac{2E}{\hbar}\left(\frac{m}{\mu}\right)^{\frac{1}{2}} = \frac{2E}{\hbar\omega_0} \\ \xi = \alpha x, \quad \psi(x) = \psi(\xi/\alpha) = \varphi(\xi); \end{aligned}\right\} \tag{5}$$

equation (3) then assumes the simpler form

$$\frac{d^2\varphi}{d\xi^2} + (\lambda - \xi^2)\varphi = 0. \tag{6}$$

We now require the eigensolutions of (6), for which $\varphi(\xi) \to 0$ as $\xi \to \pm\infty$, and the corresponding eigenvalues λ.

We shall only indicate the method of solution of this eigenvalue problem, the details of which can be found in the books by Courant–Hilbert[3], Schiff[12] or Sommerfeld[13], for example. As this eigenvalue problem is of considerable importance in many applications, we shall however state the eigensolutions and their most important properties.

To solve (6), we consider, to begin with, the asymptotic behaviour of this differential equation for large values of $|\xi|$. λ can then be neglected compared with ξ^2 and we obtain as asymptotic solutions for $|\xi| \to \infty$:

$$\varphi = e^{\pm\xi^2/2}. \tag{7}$$

Of these two solutions $\exp(+\xi^2/2)$ must be rejected as it diverges at large distances, whereas we have imposed the boundary condition that $\varphi \to 0$ at large distances.

Substituting

$$\varphi = e^{-\xi^2/2} \tag{8}$$

into (6), we find at once that it represents an exact eigenfunction corresponding to the eigenvalue

$$\lambda = 1. \tag{9}$$

Since the eigenfunction (8) possesses no nodes, this is the lowest eigenvalue, the ground-state of the harmonic oscillator according to quantum mechanics. From (5) and (9) we have for the energy E_0 of the ground-state

$$E_0 = \tfrac{1}{2}\hbar\omega_0. \tag{10}$$

Classically, we would have expected $E_0 = 0$: in the state of lowest energy the particle is at rest. According to quantum mechanics, the oscillator possesses a *zero-point energy* $E_0 = \frac{1}{2}\hbar\omega_0$. This positive zero-point energy is intimately connected with Heisenberg's uncertainty principle: the particle cannot be specified to be at rest at $x = 0$; the zero-point energy represents the energy of the fluctuations which necessarily occur.

To obtain further eigensolutions we try to find solutions of the form

$$\varphi(\xi)=e^{-\xi^2/2}\chi(\xi). \tag{11}$$

This leads to the following differential equation for $\chi(\xi)$:

$$\frac{d^2\chi}{d\xi^2}-2\xi\frac{d\chi}{d\xi}+(\lambda-1)\chi=0. \tag{12}$$

To solve equation (12), one assumes a solution in the form of a power series expansion. Either the series terminates or it does not. If it does not, it necessarily diverges. Hence the only solutions which converge as $|\xi|\rightarrow\pm\infty$ occur if the power series breaks off after a finite number of terms, *i.e.* $\chi(\xi)$ is a polynomial. The details of this analysis can be found in the books mentioned above. The eigenfunctions of (12) are known as *Hermite polynomials*,

$$\chi_n(\xi)=H_n(\xi), \quad (n=0, 1, 2, \ldots) \tag{13}$$

and the corresponding eigenvalues are

$$\lambda_n=(2n+1), \quad (n=0, 1, 2, \ldots) \tag{14}$$

i.e., from (5), the corresponding energy levels are

$$E_n=(n+\tfrac{1}{2})\hbar\omega_0, \quad (n=0, 1, 2, \ldots). \tag{15}$$

We state the most important properties of the Hermite polynomials and of the oscillator wave functions, given by equations (11) and (13):

$$\varphi_n(\xi)=e^{-\xi^2/2}H_n(\xi) \quad (n=0, 1, 2, \ldots). \tag{16}$$

(*i*) The Hermite polynomials $H_n(\xi)$ are defined explicitly by

$$H_n(\xi)=(-1)^n e^{\xi^2}\frac{d^n}{d\xi^n}(e^{-\xi^2}) \quad (n=0, 1, 2, \ldots). \tag{17}$$

(*ii*) From equations (12) and (13), $H_n(\xi)$ satisfies the differential equation

$$H_n''(\xi)-2\xi H_n'(\xi)+2nH_n(\xi)=0 \tag{18}$$

and the recurrence relation

$$H_{n+1}(\xi)-2\xi H_n(\xi)+2nH_{n-1}(\xi)=0. \tag{19}$$

(*iii*) From (17), one sees that $H_n(\xi)$ is a real polynomial in ξ (*i.e.* all the coefficients are real numbers), of order n and of parity $(-1)^n$. The first few Hermite polynomials are

$$\left.\begin{array}{ll} H_0=1 & H_1=2\xi \\ H_2=4\xi^2-2 & H_3=8\xi^3-12\xi \\ H_4=16\xi^4-48\xi^2+12 & H_5=32\xi^5-160\xi^3+120\xi. \end{array}\right\} \tag{20}$$

(*iv*) The oscillator wave functions (16), which are the solutions of our original eigenvalue problem (6), are mutually orthogonal, as we expect from the general theory of section 7(*iii*),

$$(\varphi_m, \varphi_n) = \int_{-\infty}^{\infty} \bar{\varphi}_m \varphi_n d\xi = \int_{-\infty}^{\infty} e^{-\xi^2} H_m(\xi) H_n(\xi) d\xi = 0, \quad m \neq n. \qquad (21)$$

The oscillator wave functions (16) are not yet normalized. The wave functions normed with respect to x, for which

$$(\psi_m(x), \psi_n(x)) = \int_{-\infty}^{\infty} \overline{\psi_m(x)} \psi_n(x) dx = \frac{1}{\alpha} \int_{-\infty}^{\infty} \overline{\varphi_m(\xi)} \varphi_n(\xi) d\xi = \delta_{mn} \qquad (22)$$

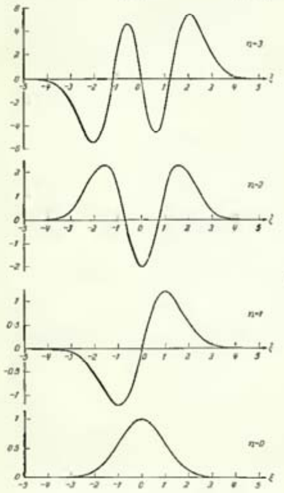

Figure 5. Harmonic oscillator wave functions $\varphi_n(\xi) = e^{-\xi^2/2} H_n(\xi)$

are
$$\psi_n(x) = \varphi_n(\xi) = \left(\frac{\alpha}{2^n n! \sqrt{\pi}}\right)^{\frac{1}{2}} e^{-\xi^2/2} H_n(\xi). \qquad (23)$$

The first few unnormalized oscillator wave functions are shown in *Figure 5.*

12. Spherically symmetric potentials in three dimensions

Finally we consider the motion of a point mass m in a field of force derived from a spherically symmetric or central potential $V(r)$, r being the distance of the particle from a fixed origin of coordinates. Atomic systems and, to a lesser degree, nuclear systems can be adequately described in terms of such central potentials which are consequently of great importance. As in classical physics, a central potential implies that the angular momentum of the system has some special properties. Since the angular momentum properties of a system are very important in quantum mechanics and as we shall investigate them in great detail in later chapters, sections 12 to 15, dealing with spherically symmetric potentials and related topics, should be studied very carefully.

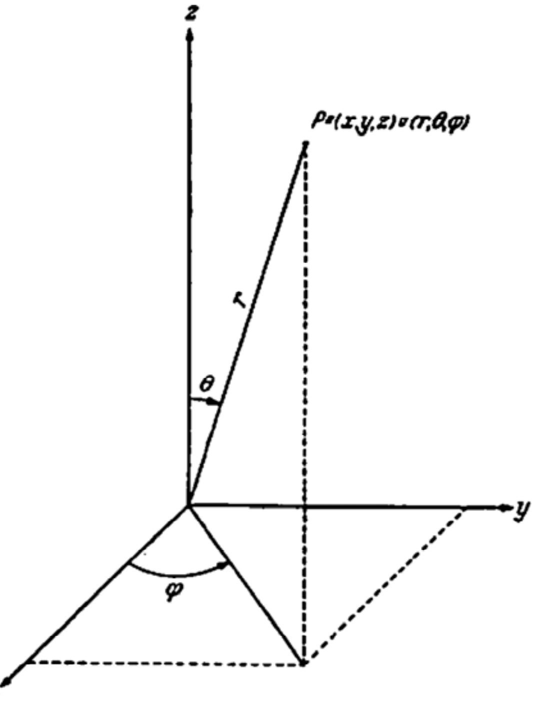

Figure 6. Spherical polar coordinates (r, θ, φ)

To study the Schrödinger equation

$$\left\{\nabla^2 + \frac{2m}{\hbar^2}[E - V(r)]\right\}\psi = 0, \tag{1}$$

it is clearly most convenient to introduce spherical polar coordinates (r, θ, φ), shown in *Figure 6*, defined by the equations:

$$\left.\begin{aligned} x &= r \sin\theta \cos\varphi \\ y &= r \sin\theta \sin\varphi \\ z &= r \cos\theta \end{aligned}\right\} \tag{2}$$

where the ranges of r, θ and φ are

$$r > 0, \quad 0 < \theta < \pi, \quad 0 < \varphi < 2\pi. \tag{3}$$

12. SPHERICALLY SYMMETRIC POTENTIALS IN THREE DIMENSIONS

In spherical polar coordinates, the Schrödinger equation (1) becomes

$$\left\{\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}\psi(r,\theta,\varphi)$$
$$+\frac{2m}{\hbar^2}[E-V(r)]\psi(r,\theta,\varphi)=0, \qquad (4)$$

where the expression in the curly bracket is the differential operator ∇^2 expressed in polar cordinates (*cf*, for example, Abraham-Becker[6], p. 43, or Hague[7], p. 112). We can separate the radial and angular variables in (4), by writing

$$\psi(r,\theta,\varphi)=R(r)S(\theta,\varphi); \qquad (5)$$

substituting this in (4) and dividing by ψ one obtains

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\frac{2m}{\hbar^2}r^2[E-V(r)]$$
$$=-\frac{1}{S}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial S}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 S}{\partial\varphi^2}\right\}. \qquad (6)$$

Since the left-hand side of this equation depends only on r and is independent of θ and φ, whereas the right-hand side depends only on θ and φ and is independent of r, it follows that the two sides must be equal to the same constant, which we shall call λ. Equation (6) then gives the pair of differential equations

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right)+\left[\frac{2m}{\hbar^2}(E-V(r))-\frac{\lambda}{r^2}\right]R=0 \qquad (7)$$

$$\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial S}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 S}{\partial\varphi^2}\right\}+\lambda S=0. \qquad (8)$$

The radial equation (7) depends on the central potential $V(r)$ which must be specified in order that (7) may be solved. Only in a few cases can the solutions be expressed in a simple closed form. These include: (*i*) a free particle, $V(r)=0$, (*ii*) a particle in a three-dimensional spherical square well, $V(r)=V_0$ for $r<r_0$, $V(r)=0$ for $r>r_0$, (*iii*) a charged particle in a Coulomb field, $V(r)=\text{const.}/r$. In general, equation (7) must be solved by some rather elaborate method, such as a series solution, numerical integration, a method of approximation, *etc.*

For these reasons, we shall not consider the radial wave equation in detail. (See, for example, the books by Schiff[12] or Sommerfeld[13, 14].) We shall only consider the asymptotic solutions for large values of r, which we shall require later. If $V(r)$ tends to zero sufficiently rapidly, as $r\to\infty$ (to be precise, if $V(r)\to 0$ more strongly than $1/r$ as $r\to\infty$), then the asymptotic solutions of (7) do not depend on $V(r)$: we can obtain them by considering the asymptotic solutions of the case $V(r)=0$. We obtain these by a method given by Sommerfeld (reference 13,

p. 22, or 14, p. 13) which is not mathematically rigorous but is simple and often useful. For large values of r equation (7) becomes

$$R'' + k^2 R = 0, \quad \left(k^2 = \frac{2mE}{\hbar^2}\right), \tag{9}$$

with two linearly independent solutions

$$R = Ae^{ikr}, \quad R = Be^{-ikr}. \tag{10}$$

One can improve the solutions (10) by taking A and B not as constants, but as slowly varying functions of r, *i.e.* we allow for terms involving the first derivative A', but neglect terms in A'', A'/r, A/r^2. Substituting (10) into (7) one then easily finds

$$A(r) = \text{const.}/r \tag{11}$$

and similarly for $B(r)$, so that we obtain as the two possible asymptotic solutions

$$R(r) = e^{\pm ikr}/r. \tag{12}$$

Any solution of (7) has an asymptotic form which is a linear combination of the two asymptotic forms (12). We see their physical significance by introducing the time-dependent factor $\exp(-i\omega t)$, equation (7.1), where $\omega = E/\hbar$. From (12) it follows that the surfaces of constant phase are

$$\pm kr - \omega t = \text{const.},$$

i.e. e^{ikr}/r represents an outgoing, diverging spherical wave and e^{-ikr}/r represents an incoming, converging spherical wave. (*Cf* Exercise III.2.)

13. Spherical harmonics

We begin by considering equation (12.8) which holds for any central potential $V(r)$. The unknown function in (12.8), $S = S(\theta, \varphi)$, depends on the polar angles θ, φ only, and can therefore be thought of as a function on the surface of the unit sphere. Equation (12.8) first arose in harmonic analysis, *i.e.* the study of solutions of the potential equation of electrostatics,

$$\nabla^2 U = 0,$$

and its solutions are called surface harmonics or *spherical harmonics*. Equation (12.8) represents an eigenvalue problem. From it we shall obtain a complete set of normed orthogonal spherical harmonics and corresponding eigenvalues; *i.e.* any sufficiently well-behaved function $f(\theta, \varphi)$, defined on the surface of the unit sphere, can be expanded in terms of these spherical harmonics.

To solve (12.8) we again separate variables; putting

$$S(\theta, \varphi) = \Theta(\theta)\Phi \tag{1}$$

it leads, by the same method as above, to the two equations

$$\frac{d^2\Phi}{d\varphi^2}+\mu\Phi=0 \tag{2}$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\left(\lambda-\frac{\mu}{\sin^2\theta}\right)\Theta=0 \tag{3}$$

where μ is a second constant, introduced in separating the angular variables θ and φ.

We must restrict the solutions $\Phi(\varphi)$ of equation (2) to be *single-valued* functions of φ as the coordinates φ and $\varphi+2\pi n$ $(n=0, \pm 1, \pm 2, \ldots)$ refer to the same point on the unit sphere and we require single-valued wave functions to lead to unambiguous physical interpretations. The required solutions are

$$\Phi_m(\varphi)=\frac{1}{\sqrt{2\pi}}e^{im\varphi} \quad (m=0, \pm 1, \pm 2, \ldots) \tag{4}$$

with

$$\mu=m^2. \tag{5}$$

We shall hereafter always assume that m has one of the values given in (4): $m=0, \pm 1, \pm 2, \ldots$ The set of solutions (4) are the complex exponential functions which we have already met in section 3, equation (3.16). As we stated there, they form a complete orthonormal set for the interval $0<\varphi<2\pi$.

To solve equation (3), with $\mu=m^2$, we introduce new variables

$$x=\cos\theta \quad \Theta(\theta)=P(x). \tag{6}$$

Equation (3) then becomes

$$\frac{d}{dx}\left[(1-x^2)\frac{dP}{dx}\right]+\left\{\lambda-\frac{m^2}{1-x^2}\right\}P=0 \tag{7}$$

and the domain of x is $-1<x<+1$ corresponding to $0<\theta<\pi$. Equation (7) is a second-order differential equation containing the unknown parameter λ, whereas m, as stated, has one of the values $m=0, \pm 1, \pm 2, \ldots$ In general, the solutions of (7) become infinite at $x=\pm 1$ and are thus physically not acceptable. Only in the special case when

$$\lambda=l(l+1) \quad l=0, 1, 2, \ldots \tag{8}$$

$$|m|<l \tag{9}$$

does (7) possess *one* solution which is everywhere finite; (the other linearly independent solution is still divergent). These solutions are called *associated Legendre functions* of order l and are defined by

$$P_l^m(x)=(1-x^2)^{|m|/2}\frac{d^{|m|}}{dx^{|m|}}P_l(x), \quad (|m|<l;\ l=0, 1, 2, \ldots) \tag{10}$$

where $P_l(x)$ are the *Legendre polynomials* of order l:

$$P_l(x)=\frac{1}{2^l l!}\frac{d^l}{dx^l}[(x^2-1)^l] \quad (l=0, 1, \ldots). \tag{11}$$

The associated Legendre functions $P_l^m(x)$ are derived from the Legendre polynomials $P_l(x)$ by a process of repeated differentiation, as shown in equation (10). For $m=0$, one has, from (10),

$$P_l^0(x)=P_l(x) \tag{12}$$

i.e. the Legendre polynomials satisfy equation (7), with $m=0$, *i.e.*

$$\frac{d}{dx}\left[(1-x^2)\frac{dP_l}{dx}\right]+l(l+1)P_l=0. \tag{13}$$

We shall not give the analysis which proves the above results. It is a rather tedious standard analysis which can be found in many books, *e.g.* those by Courant–Hilbert[3], Schiff[12] and Sommerfeld[4, 13]. We shall confine ourselves to a summary of those properties of associated Legendre functions and of spherical harmonics which are most important for quantum-mechanical applications.

(*i*) The Legendre polynomials $P_l(x)$ are real polynomials in x, of degree l and parity $(-1)^l$. The first few polynomials are

$$\left.\begin{array}{ll} P_0(x)=1 & P_1(x)=x \\ P_2(x)=\tfrac{1}{2}(3x^2-1) & P_3(x)=\tfrac{1}{2}(5x^3-3x). \end{array}\right\} \tag{14}$$

(*ii*) The associated Legendre functions $P_l^m(x)$ are of the form

$$(1-x^2)^{|m|/2}\times[\text{a real polynomial of degree } (l-|m|)],$$

and have parity $(-1)^{l-|m|}$. The first few associated Legendre functions are

$$\left.\begin{array}{l} P_1^{\pm 1}(x)=(1-x^2)^{\frac{1}{2}}, \\ P_2^{\pm 1}(x)=(1-x^2)^{\frac{1}{2}}3x, \\ P_3^{\pm 1}(x)=(1-x^2)^{\frac{1}{2}}\tfrac{3}{2}(5x^2-1), \\ P_2^{\pm 2}(x)=(1-x^2)3, \\ P_3^{\pm 2}(x)=(1-x^2)15x, \\ P_3^{\pm 3}(x)=(1-x^2)^{3/2}15. \end{array}\right\} \tag{15}$$

(*iii*) The associated Legendre functions are mutually orthogonal over the fundamental interval $-1\leqslant x\leqslant 1$,

$$\int_{-1}^{+1}P_l^m(x)P_{l'}^m(x)dx=0,\quad l'\neq l \tag{16}$$

but they are not normalized as defined in equation (10). It can be shown that

$$\int_{-1}^{+1}\{P_l^m(x)\}^2dx=\left[\frac{2}{2l+1}\,\frac{(l+|m|)!}{(l-|m|)!}\right]. \tag{17}$$

Equation (17) enables us to define normalized associated Legendre functions, apart from an arbitrary phase factor. The choice of such a factor of modulus unity does not affect the physical interpretation of the wave function (*cf* p. 28 and section 16).

In quantum mechanics it is most convenient to define the normalized associated Legendre functions by

$$(2\pi)^{\frac{1}{2}}\mathcal{N}_l^m P_l^m(x), \quad (l \geqslant |m|), \tag{18}$$

where

$$\mathcal{N}_l^m = \begin{cases} (-1)^m \left[\dfrac{2l+1}{4\pi}\,\dfrac{(l-|m|)!}{(l+|m|)!}\right]^{\frac{1}{2}}, & \text{if } m \geqslant 0, \\[2ex] \left[\dfrac{2l+1}{4\pi}\,\dfrac{(l-|m|)!}{(l+|m|)!}\right]^{\frac{1}{2}}, & \text{if } m < 0, \end{cases} \tag{19}$$

*i.e.**

$$\mathcal{N}_l^m = (-1)^{\frac{1}{2}(m+|m|)}\left[\frac{2l+1}{4\pi}\,\frac{(l-|m|)!}{(l+|m|)!}\right]^{\frac{1}{2}}. \tag{19a}$$

(*iv*) For a fixed value of m, the associated Legendre functions, defined by equations (18) and (19), form a complete normed orthogonal sequence for the fundamental interval $-1 \leqslant x \leqslant 1$.

(*v*) Combining the above results, it follows that the set of spherical harmonics, defined by

$$Y_l^m(\theta, \varphi) = \mathcal{N}_l^m P_l^m(\cos\theta)e^{im\varphi}, \quad (|m| \leqslant l, l=0, 1, \ldots) \tag{20}$$

forms a complete orthonormal system of eigenfunctions of the eigenvalue problem, equation (12.8); the corresponding eigenvalues $\lambda_l = l(l+1)$, $(l=0, 1, 2, \ldots)$ are $(2l+1)$-fold degenerate, corresponding to the $(2l+1)$ values which m can assume: $m = -l, -l+1, \ldots, l-1, l$.

A word should here be said about the, at first sight, surprising choice of phase factors, equations (19), in the definition of the spherical harmonic functions, equation (20). The reason for this choice, which is not usual in classical physics, is that spherical harmonics are closely related to the angular momentum properties of a system; in the study of these properties they occur naturally with these phase factors. For this reason, it seemed desirable to introduce the spherical harmonics in the way in which they occur in quantum mechanical applications.

From equation (19) it follows that

$$\mathcal{N}_l^{-m} = (-1)^m \mathcal{N}_l^m,$$

so that from equation (20),

$$\overline{Y_l^m(\theta, \varphi)} = (-1)^m Y_l^{-m}(\theta, \varphi).$$

One at once verifies the orthonormality of the spherical harmonics, using equations (4), (16), (17), (19) and (20):

$$\begin{aligned}(Y_{l'}^{m'}, Y_l^m) &= \int_{\substack{\text{unit}\\\text{sphere}}} d\Omega\, \overline{Y_{l'}^{m'}(\theta, \varphi)}\, Y_l^m(\theta, \varphi) \\ &= \mathcal{N}_{l'}^{m'}\mathcal{N}_l^m \int_{-1}^{+1} dx P_{l'}^{m'}(x) P_l^m(x) \int_0^{2\pi} d\varphi e^{i(m-m')\varphi} \\ &= \delta_{ll'}\delta_{mm'} \end{aligned} \tag{21}$$

---

* This form is particularly convenient in calculations as it is valid for positive and negative values of m.

where $d\Omega$ denotes an element of solid angle. That the spherical harmonics $Y_l^m(\theta, \varphi)$ are complete implies that any function $f(\theta, \varphi)$ for which

$$\int_{\substack{\text{unit}\\\text{sphere}}} d\Omega |f(\theta, \varphi)|^2$$

is finite, can be expanded in a series

$$f(\theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_l^m Y_l^m(\theta, \varphi) \tag{22}$$

and the coefficients C_l^m are determined from (21)

$$C_l^m = (Y_l^m, f) = \int_{\substack{\text{unit}\\\text{sphere}}} d\Omega \overline{Y_l^m(\theta, \varphi)} f(\theta, \varphi). \tag{23}$$

The parity of the spherical harmonics is easily found from the fact that an inversion which transforms $\mathbf{r}$ into $-\mathbf{r}$ is expressed in polar coordinates by the transformation

$$\theta \rightarrow \pi - \theta, \quad \varphi \rightarrow \varphi + \pi. \tag{24}$$

Now $\cos(\pi - \theta) = -\cos\theta$, so that

$$P_l^m(\cos[\pi - \theta]) = (-1)^{l-|m|} P_l^m(\cos\theta),$$

and

$$e^{im(\varphi+\pi)} = (-1)^{|m|} e^{im\varphi},$$

so that

$$Y_l^m(\pi - \theta, \varphi + \pi) = (-1)^l Y_l^m(\theta, \varphi) \tag{25}$$

i.e. Y_l^m has parity $(-1)^l$.

We state the first few spherical harmonics explicitly:

$$\left.\begin{aligned} &Y_0^0 = \frac{1}{\sqrt{4\pi}} \\ &Y_1^0 = \left(\frac{3}{4\pi}\right)^{\frac{1}{2}} \cos\theta, \qquad Y_1^{\pm 1} = \mp\left(\frac{3}{8\pi}\right)^{\frac{1}{2}} \sin\theta\, e^{\pm i\varphi} \\ &Y_2^0 = \left(\frac{5}{16\pi}\right)^{\frac{1}{2}} (3\cos^2\theta - 1), \ Y_2^{\pm 1} = \mp\left(\frac{15}{8\pi}\right)^{\frac{1}{2}} \cos\theta \sin\theta\, e^{\pm i\varphi}. \end{aligned}\right\} \tag{26}$$

To obtain the complete solution of the eigenvalue problem, equation (12.1), we must combine the spherical harmonic solutions (20) with the radial solutions of (12.7). For a given value of $\lambda = l(l+1)$, equation (12.7) possesses a series of radial eigenfunctions $R_{n,l}(r)$ which we shall take as normalized, corresponding to energy eigenvalues $E_{n,l}$ of the system. These energy levels may form a discrete set or a continuum depending on the nature of the potential $V(r)$. The solutions of the eigenvalue problem (12.1) are then

$$\psi_{E_{n,l},l,m}(r, \theta, \varphi) = R_{n,l}(r) Y_l^m(\theta, \varphi). \tag{27}$$

From the general discussion in Chapter II we know that these form a complete set of orthonormal eigenfunctions.

14. Angular momentum

Classically, a point particle of mass m moving in a spherically symmetric potential $V(r)$ possesses a constant angular momentum. Thus, if

$$\mathbf{p}=m\dot{\mathbf{r}},\quad \dot{\mathbf{p}}=-\operatorname{grad} V=-\frac{\partial V}{\partial r}\frac{1}{r}\mathbf{r},\quad \mathbf{l}=\mathbf{r}\wedge\mathbf{p},$$

then

$$\dot{\mathbf{l}}=\dot{\mathbf{r}}\wedge\mathbf{p}+\mathbf{r}\wedge\dot{\mathbf{p}}=m(\dot{\mathbf{r}}\wedge\dot{\mathbf{r}})-\frac{\partial V}{\partial r}\frac{1}{r}(\mathbf{r}\wedge\mathbf{r})=0,$$

i.e. the angular momentum vector **l** is constant. As there exists a close correspondence between classical observables and their quantum mechanical operators, one might expect that, quantum mechanically too, systems with central potentials have some special angular momentum properties.

Using the results of section 6, one finds for the angular momentum operator (*cf* Exercise II.1),

$$\mathbf{l}=\mathbf{r}\wedge(-i\hbar\operatorname{grad}),\tag{1a}$$

i.e.

$$l_x=-i\hbar\left(y\frac{\partial}{\partial z}-z\frac{\partial}{\partial y}\right),\ \text{etc},\tag{1b}$$

or, changing to polar coordinates,

$$\left.\begin{aligned}
l_x&=i\hbar\left(\sin\varphi\frac{\partial}{\partial\theta}+\cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right)\\
l_y&=i\hbar\left(-\cos\varphi\frac{\partial}{\partial\theta}+\cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right)\\
l_z&=-i\hbar\frac{\partial}{\partial\varphi},
\end{aligned}\right\}\tag{2}$$

and for the operator $|\mathbf{l}|$ denoting the magnitude of the angular momentum

$$|\mathbf{l}|^2=\mathbf{l}^2=l_x^2+l_y^2+l_z^2=-\hbar^2\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}.\tag{3}$$

Consider a wave function

$$\psi_l^m(r,\theta,\varphi)=R(r)Y_l^m(\theta,\varphi),\quad(|m|<l,\ l=0,1,\ldots),\tag{4}$$

where the angular dependence is to be given by the spherical harmonic $Y_l^m(\theta,\varphi)$, but the radial wave function, $R(r)$, in which we are not interested, is not specified any further. From equations (2) and

(13.20) it follows that

$$l_z\psi_l^m = (\hbar m)\psi_l^m, \quad (|m| \leqslant l,\ l=0,\ 1,\ \ldots), \tag{5}$$

i.e. ψ_l^m is an eigenfunction of the operator l_z, representing the z-component of angular momentum, with eigenvalue $(\hbar m)$. m is called the *magnetic quantum number* of the system, as it is of particular importance in studying the effect of magnetic fields on atomic spectra. From equations (12.8), (13.8), (13.20) and (3) one further has, that

$$\mathbf{l}^2\psi_l^m = \hbar^2 l(l+1)\psi_l^m, \quad (|m| \leqslant l,\ l=0,\ 1\ \ldots), \tag{6}$$

i.e. ψ_l^m is an eigenfunction of the operator $\mathbf{l}^2$ with eigenvalue $\hbar^2 l(l+1)$. l is called the (orbital) *angular momentum quantum number* of the particle.

From the physical interpretation of eigenfunctions and eigenvalues, given in section 6, this implies that a particle described by a wave function ψ_l^m has an exact value of angular momentum $|\mathbf{l}|$ equal to $\hbar\sqrt{l(l+1)}$, and the z-component of angular momentum has the exact value $\hbar m$. Thus in the state ψ_l^m, which is a simultaneous eigenstate of the operators $\mathbf{l}^2$ and l_z, the exact values of the corresponding observables $\mathbf{l}^2$ and l_z can be specified.

On the other hand, ψ_l^m is not an eigenfunction of l_x or l_y as one easily verifies from equations (2) for any particular spherical harmonic*. In Chapter IV we shall see that this implies that in the state ψ_l^m the x- and y-components of angular momentum do not have sharply defined values. This means that if we know that the system is in the state ψ_l^m (*i.e.* we measured $\mathbf{l}^2$ and l_z and found the values $\hbar^2 l(l+1)$ and $\hbar m$ respectively), then measuring l_x or l_y does not produce the same result if we repeat the experiment several times (each time starting from the same state ψ_l^m), but there are several different possible results which have definite probabilities of occurring. These questions will be examined much more carefully in the next chapter.

One other point should be mentioned. At first sight, it appears as though the z-direction is in some way distinguished or preferred, in spite of the fact that the original problem is specified by a spherically symmetric potential $V(r)$. But this is not the case. For a given value of the angular momentum $\hbar\sqrt{\{l(l+1)\}}$ there exist $(2l+1)$ linearly independent eigenfunctions ψ_l^m, $m=-l,\ \ldots l$; in general the wave function will be a linear combination of these:

$$\psi_l = \sum_{m=-l}^{l} c_m\psi_l^m,$$

but we are not at liberty to choose the linear combination arbitrarily;

---

* The case $l=0$ represents an exception. Equation (6) now states that the total angular momentum $|\mathbf{l}|$ is exactly zero, and one would expect any component of $\mathbf{l}$ also to be exactly zero. In fact, since $\psi_0^0 = R(r)/\sqrt{(4\pi)}$, it follows at once from (2) that ψ_0^0 is a simultaneous eigenfunction of all three components of angular momentum, with the common eigenvalue zero.

it is determined by the physical conditions. By a suitable choice of the coefficients c_m we can make ψ_l represent a state with a definite component of angular momentum, $\hbar m$, in *any* prescribed direction. For a spherically symmetric system we must average over all possible values of c_m, *i.e.* average over all possible orientations. The spherical harmonics, *i.e.* the ψ_l^m, are themselves already the correct linear combinations of wave functions if there exists a preferred direction in the the system which is taken as z-axis; for example, if a magnetic field is applied in the z-direction.

The properties of angular momentum which we have so far considered have all referred to one instant of time. In Chapter IV we shall see that theorems analogous to the classical conservation theorems hold in quantum theory also.

EXERCISES III

III.1. Suppose that the Hamiltonian function H of a system is a function of n position and momentum vectors $\mathbf{r}_1, \ldots, \mathbf{r}_n, \mathbf{p}_1, \ldots \mathbf{p}_n$,

$$H = H(\mathbf{r}_1, \ldots \mathbf{r}_n, \mathbf{p}_1, \ldots \mathbf{p}_n)$$

and that it is invariant under inversion of coodinates *i.e.*

$$H(-\mathbf{r}_1, \ldots -\mathbf{r}_n, -\mathbf{p}_1, \ldots -\mathbf{p}_n) = H(\mathbf{r}_1, \ldots \mathbf{r}_n, \mathbf{p}_1, \ldots \mathbf{p}_n). \quad \text{(III.1)}$$

If $\psi(\mathbf{r}_1, \ldots \mathbf{r}_n)$ is an eigenfunction of H belonging to the *non-degenerate* eigenvalue E, show that ψ has a definite parity, *i.e.*

$$\psi(-\mathbf{r}_1, \ldots -\mathbf{r}_n) = \pm\psi(\mathbf{r}_1, \ldots \mathbf{r}_n).$$

III.2. Show that the wave functions

$$\frac{1}{r}e^{\pm ikr}$$

represent spherical outgoing and incoming waves, respectively, by considering the corresponding probability currents.

III.3. The Schrödinger equation of two interacting particles of mass m_1 and m_2 respectively is

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}_1, \mathbf{r}_2, t) = \left[-\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(\mathbf{r}_1 - \mathbf{r}_2)\right]\psi(\mathbf{r}_1, \mathbf{r}_2, t), \quad \text{(III.3)}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position coordinates of the particles and

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}, \quad (i = 1, 2).$$

Show that if one introduces the centre of gravity coordinate,

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2},$$

and the relative coordinate $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ of the two particles, then the Schrödinger equation (III.3) is separable. Interpret the two resulting equations.

III.4. In section 8 we considered a plane wave incident on a potential jump, equation (8.2).

Consider the same problem for a potential barrier of finite thickness:

$$\begin{aligned} V(x) &= 0, && x<0,\\ V(x) &= V_0, \quad (V_0>0), && 0<x<a,\\ V(x) &= 0, && x>a, \end{aligned}$$

and obtain the reflection and transmission coefficients for both the cases $E>V_0$ and $0<E<V_0$.

III.5. Obtain the eigenfunctions and eigenvalues of the eigenvalue problem

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right]\psi(x)=E\psi(x),$$

where

$$\begin{aligned} V(x) &= 0, && |x|<a,\\ V(x) &= V_0, \ (V_0>0), && |x|>a, \end{aligned}$$

which lie in the interval $0<E<V_0$. (In this energy range, this problem has a discrete spectrum of eigenvalues.)

(Hint: The solution of this eigenvalue problem can be simplified in two respects: (*i*) by using the fact that the solutions must be even or odd (*cf* Example III.1), one need only solve the problem for $x>0$; (*ii*) instead of matching the wave function and its derivative at $x=a$, we need only match its *logarithmic derivative*; for

$$\psi_1(a)=\psi_2(a), \quad \psi_1'(a)=\psi_2'(a)$$

implies

$$\frac{\psi_1'(a)}{\psi_1(a)}=\frac{\psi_2'(a)}{\psi_2(a)}$$

i.e.

$$\left[\frac{d}{dx}\ln\psi_1(x)\right]_{x=a}=\left[\frac{d}{dx}\ln\psi_2(x)\right]_{x=a}.$$

This procedure of matching logarithmic derivatives is generally applicable and often advantageous.)

III.6. Obtain the Schrödinger wave equation of the hydrogen atom. Obtain the eigenfunction corresponding to the lowest energy, which is spherically symmetric. In this ground-state, what is the most probable distance between the electron and the proton?

Obtain the ground-state wave function of the helium atom, neglecting the Coulomb repulsion between the two electrons.

IV

GENERAL PRINCIPLES OF QUANTUM MECHANICS

In Chapter II we introduced Schrödinger's wave equation and dealt with its more important mathematical consequences and, to a lesser extent, with its physical interpretation. Our main object in this chapter is to investigate systematically the physical interpretation, *i.e.* the relations between the mathematical symbols and their laws of operation on the one hand, and the physical observables and their measurement on the other. At the same time we shall extend the mathematical formalism to admit much more general questions than we have considered so far. It will turn out that all questions concerning the interpretation of quantum mechanics can be summarized in five basic rules or interpretative postulates, which we shall state and discuss in section 16. We shall try to make these axioms appear physically reasonable, and show that they are often natural generalizations of concepts we have already met. These axioms will merely serve as simple and concise rules for the interpretation of quantum theory. We shall not discuss the difficult question of whether these axioms are the minimum number of necessary axioms, or examine their reasonableness in terms of highly idealized measurements. Such analysis is of course of great importance, but falls entirely outside the scope of this book.

On the basis of these postulates we shall, in the remainder of the chapter, consider general questions of interpretation, such as the measurement of several observables, the uncertainty principle, *etc.*

15. Résumé of classical mechanics*

Before considering the general principles of quantum mechanics, we shall briefly restate the laws of classical mechanics and their consequences.

The configuration of a system of N point masses $m_1, \ldots m_N$ is completely given by the $3N$ Cartesian coordinates $(x_1, y_1, z_1), (x_2, y_2, z_2) \ldots (x_N, y_N, z_N)$ of the masses. Frequently it is not convenient to restrict oneself to these Cartesian coordinates, but to specify the configuration of the system in terms of other parameters, called *generalized coordinates*, *e.g.* we sometimes prefer to work in polar coordinates (r, θ, φ). This is particularly the case if the system is subject to constraints, so that the $3N$ Cartesian coordinates are not independent variables. If n is the least number of parameters necessary to specify the configuration of the system, then the system is said to have *n degrees of freedom.* (Thus a point mass has 3 degrees of freedom, a rigid body has 6 degrees of freedom.) The configuration of a system with n degrees of freedom is fully specified by n generalized position coordinates $q_1, q_2, \ldots q_n$, and we express all quantities, in particular the Cartesian coordinates (x_a, y_a, z_a), in terms of these generalized coordinates. The kinetic energy T and the potential energy V of the system become

---

* Section 15 will not be required subsequently, but has been included to bring out the essential contrast between classical and quantum mechanics. For a more detailed treatment see, for example, M. Planck, *General Mechanics*, London: Macmillan, 1933.

$$T = \tfrac{1}{2}\sum_{a=1}^{N} m_a\left\{\left(\frac{dx_a}{dt}\right)^2 + \left(\frac{dy_a}{dt}\right)^2 + \left(\frac{dz_a}{dt}\right)^2\right\} = T(q_i, \dot{q}_i), \tag{1}$$

$$V = V(q_i, \dot{q}_i) \tag{2}$$

where we have abbreviated the set of coordinates $(q_1, q_2, \ldots q_n)$ to (q_i) and the dot denotes differentiation with respect to the time: $\dot{q}_i = dq_i/dt$. From (1), we define the *generalized momenta* by

$$p_i = \frac{\partial}{\partial \dot{q}_i}(T - V), \quad (i = 1, 2, \ldots n), \qquad p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{3}$$

analogous to their definition in Cartesian coordinates for systems for which, as is generally the case, the potential energy V is velocity-independent (*i.e.* $\partial V/\partial \dot{q}_i = 0$, $i = 1, 2, \ldots n$),

$$p_{x_a} = m_a \dot{x}_a = \frac{\partial T}{\partial \dot{x}_a}, \quad (a = 1, \ldots N),$$

and similarly for the other components. We now define the *Hamiltonian function H* of the system as its energy function expressed in terms of the $2n$ generalized position and momentum coordinates,

$$H = H(q_i, p_i). \tag{4}$$

In terms of this Hamiltonian function, the equations of motion, *Hamilton's equations*, become

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \dot{q}_i = \frac{\partial H}{\partial p_i}, \quad (i = 1, \ldots n). \tag{5}$$

Hamilton's equations are the generalization of Newton's equations of dynamics, and for Cartesian coordinates they reduce to these (*cf* Exercises IV, 1 and 2). Hamilton's equations form a system of $2n$ partial differential equations of the first order for the $2n$ unknown functions $q_i(t)$, $p_i(t)$, $(i = 1, \ldots n)$; given the values of the $2n$ functions at some instant of time t_0, Hamilton's equations possess a unique solution, *i.e.* they completely specify the position and momentum coordinates at all other times. In classical mechanics all dynamical quantities, such as energy, angular momentum, *etc*, are well-defined functions of the generalized coordinates. Hence, given the initial values of these at

$=t_0$, the subsequent behaviour and properties of the system are completely determined by Hamilton's equations. The fundamental laws of classical mechanics are completely deterministic. Classical statistical mechanics is merely the result of averaging, when one does not know the initial values of all the $2n$ coordinates of a macroscopic system.

16. Interpretative postulates of quantum mechanics

We turn to the quantum mechanical description of a system with n degrees of freedom which is classically specified by n generalized position coordinates $q_1, q_2, \ldots q_n$.

In quantum mechanics such a system is specified at a fixed instant of time by a wave function $\psi(q_1, \ldots q_n)$ whose norm is unity, *i.e.*

$$(\psi, \psi) = \int |\psi(q_1, \ldots q_n)|^2 dq_1 \ldots dq_n = 1, \qquad (1)$$

the integration being over the fundamental domain, *i.e.* the accessible values of the coordinates $q_1, \ldots, q_n$. The set of all such wave functions with unit norm span a function space or Hilbert space $\mathfrak{H}$, in the way which we discussed in detail in Chapter I. Every possible state of the system is represented by a function in this Hilbert space, and, conversely, every vector in this space represents a possible state of the system. In addition to depending on the coordinates $q_1, \ldots q_n$, the wave function depends also on the time t, but the dependence on the qs and on t are essentially different. The Hilbert space $\mathfrak{H}$ is formed with respect to the spatial coordinates $q_1, \ldots q_n$ only, *e.g.* the inner products are formed with respect to the q's only, and one wave function $\psi(q_1, \ldots q_n)$ states its complete spatial dependence. On the other hand the states of the system at different instants of time $t_1, t_2, \ldots$ are given by different wave functions $\psi_{t_1}(q_1, \ldots q_n)$, $\psi_{t_2}(q_1, \ldots q_n)$, $\ldots$ of the Hilbert space. This mode of description which treats the time t as a parameter on an essentially different footing from the position coordinates results from the non-relativistic way in which we are formulating quantum mechanics.

The normalization of the state vectors ψ_t is closely connected with the probability interpretation of the wave function. This normalization, however, still leaves undetermined a factor of modulus unity in the wave function. If α is real, then $\|e^{i\alpha}\psi_t\| = \|\psi_t\|$. We shall see that such a factor of modulus unity in the wave function does not affect the predictions of observable physical properties, *i.e.* of experiments. Thus the choice of a phase factor $e^{i\alpha}$ in the wave function is arbitrary, but in a quantum mechanical calculation involving different wave functions one must of course fix these phase factors in a consistent manner, as they determine the interference between these waves.

It is one of the most basic assumptions of quantum theory that a complete knowledge of the wave function represents the maximum information of the system that is possible. Any information that one can give about the system can be derived from its wave function, and no experiments can give more information than can be represented by a wave function. In this statement is already implied, as we shall see later, the inherent inability of quantum mechanics to specify a system as fully as is the case in classical theory. We shall see presently how one derives information about the system from its wave function. Here we only stress the severe limitation that there exists no way of obtaining additional information about the system beyond that which can be derived from the wave function. We summarize these results in:

> *Postulate I.* A system with n degrees of freedom is completely specified by a normed state vector $\psi_t(q_1, \ldots q_n)$ which contains an arbitrary factor of modulus unity. All possible information about the system can be derived from this wave function.

The state of a system, specified by a vector in Hilbert space, represents a theoretical abstraction. One cannot measure the state directly in any way. What one does do is to measure certain physical quantities such as energies, momenta, *etc.* Following Dirac, we shall call these observables. From these observations one then has to infer the state of the system.

Observables are represented in quantum mechanics by operators. Mathematically, an operator A defines a transformation in the function space describing the system, which transforms one wave function ψ into another wave function χ:

$$\chi = A\psi.$$

We can think of A as defining a correspondence between the two wave functions, as a rule for associating a vector χ with the vector ψ. In Chapter II we had some examples of operators. We defined the operator x_{op} as 'multiplication by x' and x_{op} operating on $\psi(x)$ gives $\chi(x) = x\psi(x)$. Or we can define the operator A as 'differentiation with respect to x'; in that case

$$\chi(x) = A\psi(x) = \frac{\partial}{\partial x}\psi(x).$$

The rule defining the correspondence between the object function ψ and the image function χ need not of course be as simple as in these examples, but it is always useful to interpret it geometrically: ψ and χ are two vectors in Hilbert space, and A transforms, *i.e.* rotates and magnifies, ψ to become identical with χ.

The operators of quantum mechanics are linear, *i.e.* for any two complex numbers c_1 and c_2

$$A(c_1\psi_1 + c_2\psi_2) = c_1 A\psi_1 + c_2 A\psi_2.$$

This property of being linear is a most important one which is constantly used.

16. INTERPRETATIVE POSTULATES OF QUANTUM MECHANICS

In Chapter II we considered the operators $\mathbf{r}_{op}$ and $\mathbf{p}_{op}$ corresponding to the classical observables position vector $\mathbf{r}$ and momentum vector $\mathbf{p}$. We saw that

$$\mathbf{r}_{op} \equiv \mathbf{r}, \quad \mathbf{p}_{op} \equiv -i\hbar \text{ grad},$$

so that

$$\chi(\mathbf{r}) = \mathbf{r}_{op}\psi(\mathbf{r}) = \mathbf{r}\psi(\mathbf{r})$$

and

$$\chi(\mathbf{r}) = \mathbf{p}_{op}\psi(\mathbf{r}) = -i\hbar \text{ grad } \psi(\mathbf{r}),$$

and that the classical observable $F(x_1, \dots z_N; p_{x_1}, \dots p_{z_N})$ represented by the operator $F(x_1, \dots z_N; -i\hbar\partial/\partial x_1, \dots -i\hbar\partial/\partial z_N)$. One can show that if one uses generalized coordinates q_i, $(i=1, \dots n)$, the operators representing the generalized momenta p_i, equation (15.3), are $-i\hbar\partial/\partial q_i$, and the classical observable $F(q_1, \dots q_n; p_1, \dots p_n)$ is represented by the operator $F(q_1, \dots q_n; -i\hbar\partial/\partial q_1, \dots -i\hbar\partial/\partial q_n)$. This is a natural generalization of our results for Cartesian coordinates.

We shall say that φ is an eigenvector of A with a as eigenvalue, if

$$A\varphi = a\varphi.$$

We shall assume that operators corresponding to observables possess a complete set of eigenvectors $\varphi_1, \varphi_2, \dots \varphi_m \dots$ and corresponding discrete eigenvalues $a_1, a_2, \dots a_m \dots$ ordered so that $Re\, a_1 < Re\, a_2 < \dots < Re\, a_n \dots (Re\, a_n = \text{real part of } a_n)$; then

$$A\varphi_n = a_n\varphi_n \tag{2}$$

and any state ψ with finite norm can be expanded in a linear superposition of these eigenstates of A

$$\psi = \sum_{n=1}^{\infty} c_n\varphi_n. \tag{3}$$

In addition to these operators with completely discrete spectra, one also requires operators with continuous spectra to represent observables. For example, the position vector of a particle has a continuous spectrum. In section 17 we shall give the necessary generalizations to deal with continuous spectra. But the physical interpretation is not much affected by this change. Consequently we shall restrict ourselves to operators with discrete spectra to keep the mathematics as simple as possible*.

To complete the definition of operators which represent observables, we impose the restriction that they are *Hermitian* or *self-adjoint*. A is a Hermitian operator if for any vector ψ, the scalar product $(\psi, A\psi)$ is real, *i.e.* if

$$(\psi, A\psi) = (A\psi, \psi). \tag{4}$$

---

* Where the nature of the spectrum is not important, we shall also use operators with continuous spectra as illustrative examples.

We shall see below that $(\psi, A\psi)$ represents the expectation value of the observable A for a system in a state ψ, and this is the reason we restrict ourselves to Hermitian operators. We can see this differently by taking one of the eigenfunctions of A, say φ_n, for ψ in equation (4). It then becomes

$$a_n = \bar{a}_n, \tag{5}$$

i.e. the eigenvalues of a Hermitian operator are real, and we know from Chapter II that the eigenvalue a_n is the value of the observable A in the state φ_n.

Since equation (4) is to hold for any vector ψ, take $\psi = c_1\chi_1 + c_2\chi_2$, where c_1 and c_2 are arbitrary complex numbers and χ_1 and χ_2 are arbitrary vectors. Then

$c_1 \chi_1 + c_2$ [illegible]

$$(\psi, A\psi) = |c_1|^2(\chi_1, A\chi_1) + |c_2|^2(\chi_2, A\chi_2) + \{c_1\bar{c}_2(\chi_2, A\chi_1) + \bar{c}_1c_2(\chi_1, A\chi_2)\}$$

and this is real for all values of c_1 and c_2 if, and only if,

$$\underline{(\chi_2, A\chi_1) = (A\chi_2, \chi_1).} \tag{6}$$

Equation (6) represents a very important property of a Hermitian operator and may be taken for its definition.

Examples of self-adjoint operators are the position and momentum coordinates. In the one-dimensional case one has

$$(\chi, x\psi) = \int_{-\infty}^{\infty} \bar{\chi}x\psi dx = \int_{-\infty}^{\infty} \overline{(x\chi)}\psi dx = (x\chi, \psi),$$

since x is real, and

$$(\chi, p_x\psi) = \int_{-\infty}^{\infty} \bar{\chi}\left(-i\hbar\frac{\partial}{\partial x}\psi\right)dx = \left[-i\hbar\bar{\chi}\psi\right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty}\left(i\hbar\frac{\partial}{\partial x}\bar{\chi}\right)\psi dx = (p_x\chi, \psi),$$

since $\left[-i\hbar\bar{\chi}\psi\right]_{-\infty}^{\infty}$ vanishes, as ψ and χ tend to zero when $x \to \pm\infty$.

One can easily prove that eigenfunctions belonging to different eigenvalues of a Hermitian operator are orthogonal; for

$$0 = (\varphi_m, A\varphi_n) - (A\varphi_m, \varphi_n) = (a_n - a_m)(\varphi_m, \varphi_n),$$

and if $a_n \neq a_m$, then,

$$(\varphi_m, \varphi_n) = 0.$$

If $a_n = a_m$, *i.e.* for degenerate eigenfunctions, we can always choose mutually orthogonal linear combinations of eigenfunctions, using the Schmidt orthogonalization process (section 3). For the particular case of the energy operator we already proved the reality of the eigenvalues and the orthogonality of the eigenfunctions in section 7. We state our general results about observables and operators in:

> *Postulate II.* To every observable corresponds a Hermitian operator with a complete set of orthonormal eigenfunctions.

So far we have only considered one instant of time. The temporal variation of a system is given by

Postulate III. For every system there exists a Hermitian operator H, the Hamiltonian operator, which determines the time variation of the system through the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t}=H\psi, \tag{7}$$

provided the system is not disturbed; measurement represents such a disturbance.

We have already discussed the time-dependent Schrödinger equation in Chapter II. It is the equation of motion of quantum mechanics. Like the equations of motion of classical mechanics it is completely deterministic: given the wave function at some time $t=t_0$, equation (7) uniquely determines the wave function ψ_t at any other time t. Equation (7) only applies as long as the Hamiltonian H describes the system as a whole, *i.e.* the Hamiltonian must include any 'external' interactions with the system. If, for example, we consider an atom in an external magnetic field, the Hamiltonian must contain a term representing the interaction of the atom with the magnetic field. In other words, like the classical equations, equation (7) only holds for closed systems, *i.e.* for systems not disturbed from outside. According to classical physics, we can carry out measurements on the system which interfere to an arbitrarily small extent with it. According to quantum theory this is not the case. The interactions between the system and the measuring apparatus cannot be made arbitrarily weak but there exists a definite finite lower limit. As a result, a measurement in general represents a disturbance of the system and its temporal development is no longer given by equation (7). If ψ_0 is the wave function at the initial time $t_0=0$, we can integrate equation (7) to give the solution

$$\psi_t=\exp\left\{-\frac{i}{\hbar}\int_0^t Hdt\right\}\psi_0. \tag{8}$$

We have not yet defined what we mean by the rather complicated operator

$$S_t=\exp\left\{-\frac{i}{\hbar}\int_0^t Hdt\right\} \tag{9}$$

operating on ψ_0 and we shall do this in section 18 dealing with functions of operators. But it follows from equation (7) that, geometrically speaking, for a given initial vector ψ_0, ψ_t traces out a completely defined trajectory in function space. The operator S_t relates the vectors associated with the initial point $t=0$ and an arbitrary point t on the trajectory. In section 18 we shall also show that S_t has the property that

$$(S_t\psi_0, S_t\psi_0)=(\psi_0, \psi_0) \tag{10}$$

i.e.

$$(\psi_t, \psi_t)=(\psi_0, \psi_0); \tag{11}$$

in other words, the normalization of the wave function persists throughout time, a result which we had already established in section 6.

The general solution, equation (8), of the Schrödinger equation simplifies if the Hamiltonian function H does not depend on the time explicitly ($\partial H/\partial t=0$); it then reduces to

$$\psi_t=T_t\psi_0, \tag{12}$$

with

$$T_t=\mathrm{e}^{-iHt/\hbar}. \tag{13}$$

We have introduced wave functions as representing states and operators as representing the physical observables of systems. We must now see how these quantities are related to actual measurements. Given a system in a definite state, described by a given wave function, how can we predict its physical properties, which are to be compared with experiments? Conversely, how do we incorporate information about a system, gained as a result of measurements, into the wave function? How, for example, do we determine the initial state ψ_0 of a system which we must know to be able to integrate the wave equation (7)? For this analysis, and in the rest of this chapter, we shall use Gothic letters to denote observables when we want to stress the distinction between the observables and the corresponding operators.

In the first place, we consider the measurement of a single observable $\mathfrak{A}$. Only in sections 19 and 20 shall we consider the measurement of several observables. We shall again assume that the operator A, corresponding to this observable $\mathfrak{A}$, has a complete normed orthogonal sequence of eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots$ and corresponding eigenvalues $a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant \ldots$ We have stated already in section 6 that if the system is in an eigenstate φ_n of A, then, measuring $\mathfrak{A}$ gives the certain result a_n. In general, the system will be in a state ψ (with $\|\psi\|=1$), which is not an eigenstate of A. But as we know, ψ can always be expanded in the form

$$\psi=\sum_{n=1}^{\infty}c_n\varphi_n \tag{14}$$

where $c_n=(\varphi_n, \psi)$, *i.e.* the state ψ is a linear superposition of the different eigenstates $\varphi_1, \varphi_2, \ldots$, in each of which $\mathfrak{A}$ would have a perfectly definite value $a_1, a_2, \ldots$ respectively. Measuring $\mathfrak{A}$ on a system in a state ψ no longer leads to a definite result; rather there exists a probability distribution to find any one of the values $a_1, a_2, \ldots$ This probability distribution is given by

Postulate IV. If $\mathfrak{A}$ is measured on a system in a state ψ, equation (14), the probability of finding for $\mathfrak{A}$ a value a lying in the interval $a' < a < a''$ is

$$P(a', a'') = \sum |c_n|^2, \tag{15}$$

the summation being over *only* those values of n for which a_n lies in the interval $a' < a < a''$.

Postulate IV contains the basic rule for predicting the result of measurements. Its most important implication is that, in contradistinction to classical mechanics, a system with the maximal specification of its state, namely with a given wave function, still shows dispersion: measuring $\mathfrak{A}$ does not necessarily lead to one definite result. Rather, there exists only a probability distribution, which can be calculated, of finding a particular result. This means that if we measured $\mathfrak{A}$ on a large number of systems, all in the state ψ, then the different possible results $a_1, a_2, \ldots$ would occur with a frequency given by Postulate IV. Thus the diffraction pattern of an electron scattering experiment corresponds to the probabilities of the electron being scattered into different final states with different momentum vectors. One can obtain such a pattern 'instantaneously' by scattering a shower of electrons, or by scattering successive individual electrons and recording these events on a photographic plate.

We obtain a definite result if the system is in an eigenstate φ_n of A before the measurement. In that case there is a probability 1 of observing the value a_n and 0 of getting any other result, as follows at once from (15). Conversely, if we know that the system is in a state ψ such that on measuring $\mathfrak{A}$ one certainly finds a, then a must be an eigenvalue of A and ψ a corresponding eigenfunction: $A\psi = a\psi$, i.e. ψ is of the form $\psi = \Sigma c_n \varphi_n$, the summation being only over those values of n for which $a_n = a$. From (15) it follows that one can carry out measurements of absolute accuracy in quantum mechanics, provided A has a discrete spectrum. For if the interval (a', a'') is sufficiently small so that it contains one eigenvalue a_n only, $a' < a_n < a''$, then the result '$\mathfrak{A}$ lies in (a', a'')' implies that $\mathfrak{A}$ has the exact value a_n. From equation (15) and the fact that $\|\psi\| = 1$, it follows that

$$P(-\infty, \infty) = \sum_{n=1}^{\infty} |c_n|^2 = 1, \tag{16}$$

i.e. the probabilities in equation (15) are correctly normalized.

Since $\mathfrak{A}$ does not necessarily have a definite value, we are interested in its statistics. We define its expectation value $\langle A \rangle$ in the state ψ by

$$\langle A \rangle = \sum_{n=1}^{\infty} |c_n|^2 a_n. \tag{17}$$

$\langle A\rangle$ represents the average value which one finds on repeatedly measuring $\mathfrak{A}$, always starting with a system in the state ψ. An alternative way of writing (17) is

$$\langle A\rangle = (\psi, A\psi) \tag{18}$$

and we see from equations (17) or (18) that $\langle A\rangle$ is real, as required. If $\psi = \varphi_n$, equation (18) gives $\langle A\rangle = a_n$.

In addition to the expectation value we are interested in the mean square deviation ΔA which is a measure of the dispersion around the mean value $\langle A\rangle$. We define $(\Delta A)^2$ as the expectation value of $[A - \langle A\rangle]^2$ in the state ψ,

$$\begin{aligned}(\Delta A)^2 &= \langle [A - \langle A\rangle]^2\rangle = (\psi, [A - \langle A\rangle]^2\psi)\\ &= (\psi, [A^2 - 2\langle A\rangle A + \langle A\rangle^2]\psi) = (\psi, A^2\psi) - \langle A\rangle^2\end{aligned}$$

so that

$$(\Delta A)^2 = \|[A - \langle A\rangle]\psi\|^2 = \langle A^2\rangle - \langle A\rangle^2, \tag{19}$$

where, on account of equation (4),

$$\langle A^2\rangle = (\psi, A^2\psi) = (\psi, A(A\psi)) = (A\psi, A\psi) = \|A\psi\|^2.$$

The condition that the mean square deviation vanishes, $\Delta A = 0$, implies $[A - \langle A\rangle]\psi = 0$, *i.e.*

$$A\psi = \langle A\rangle\psi,$$

i.e. $\langle A\rangle$ is an eigenvalue of A belonging to the eigenfunction ψ. This shows that $\Delta A = 0$ is a necessary and sufficient condition for the system to be in an eigenstate of A.

We illustrate the above results by considering the measurement of the z-component l_z of the orbital angular momentum of an atomic electron, treated in section 14. We shall consider two distinct cases in which the wave functions of the electron are given by equation (14.4),

$$\psi_l^m(\mathbf{r}) = R(r)Y_l^m(\theta, \varphi), \tag{14.4}$$

and by

$$\psi(\mathbf{r}) = R(r)\frac{1}{\sqrt{2l+1}}\sum_{m=-l} Y_l^m(\theta, \varphi) \tag{20}$$

respectively, where we assume

$$\int_0^\infty |R(r)|^2 r^2 dr = 1$$

so that ψ_l^m and ψ are normed to unity. Both ψ_l^m and ψ are eigenstates of the operator $\mathbf{l}^2$ giving the square of the magnitude of the total orbital angular momentum of the electron [equation (14.6)], hence in both states ψ_l^m and ψ the angular momentum of the electron has the value $\hbar\sqrt{\{l(l+1)\}}$.

ψ_l^m is an eigenfunction of l_z with eigenvalue $\hbar m$ [equation (14.5)]:

$$l_z\psi_l^m = \hbar m \psi_l^m. \tag{14.5}$$

In this case we have not only $\langle l_z\rangle = \hbar m$, but also $\Delta l_z = 0$: measuring l_z leads to the exact result $\hbar m$.

ψ is not an eigenfunction of l_z. One easily finds

$$\langle l_z\rangle = 0 \tag{21}$$

and (*cf* Exercise IV.4)

$$(\Delta l_z)^2 = \langle l_z^2\rangle = \tfrac{1}{3}\hbar^2 l(l+1). \tag{22}$$

For the probability of finding $\hbar m$ $(m = -l, \ldots +l)$ for the value of l_z in the state ψ, one has, from equation (15),

$$P(\hbar m) = \frac{1}{2l+1} \quad (m = -l, \ldots l). \tag{23}$$

We can interpret ψ_l^m, somewhat classically, as a state in which the angular momentum vector $\mathbf{l}$, of magnitude $\hbar\sqrt{\{l(l+1)\}}$, makes a definite angle α with the z-axis, where

$$\cos\alpha = \frac{m}{\sqrt{l(l+1)}}. \tag{24}$$

We might have expected the maximum value of l_z to be $\hbar\sqrt{\{l(l+1)\}}$, corresponding to $\alpha = 0$, *i.e.* the momentum vector $\mathbf{l}$ pointing in the direction of the z-axis. That the maximum value of l_z is $\hbar l$, corresponding to $\cos\alpha = \sqrt{\{l/(l+1)\}}$, is closely related to the uncertainty principle. According to this, the exact values of l_x and l_y cannot be specified simultaneously with l_z. $\alpha = 0$, however, means that l_x and l_y have the exact values zero. We shall return to this point in section 20.

On the other hand, we can interpret ψ as a state in which the angular momentum vector $\mathbf{l}$ has random orientation. It describes a spherically symmetric system in which no direction is preferred. This follows from equations (21) to (23); equations (21) and (22), in particular, follow from the fact that for a spherically symmetric system $\langle l_z\rangle = 0$, and

$$\langle l_x^2\rangle = \langle l_y^2\rangle = \langle l_z^2\rangle = \tfrac{1}{3}\hbar^2 l(l+1),$$

since $\mathbf{l}^2$ has the value $\hbar^2 l(l+1)$. Equation (23) expresses the fact that all values of m, *i.e.* all orientations of the vector $\mathbf{l}$ are equally probable; they are independent of m.

We come to the converse question of how information obtained about a system by measurements is incorporated in the wave function.

Suppose again, that the observable $\mathfrak{A}$ is measured on a system in a state ψ, and the value a (an eigenvalue of A, according to Postulate IV) is observed for $\mathfrak{A}$. If this statement is to mean anything it must imply that if one repeats the measurement of $\mathfrak{A}$ sufficiently quickly one necessarily again finds the same value a. In general, this only holds for a sufficiently short interval of time between the two measurements. For if the interval is too long the system will have changed appreciably merely as a consequence of the time variation of the system according to Schrödinger's equation (7), and the value of $\mathfrak{A}$ will also have changed. Only if $\mathfrak{A}$ is a constant of the motion (*cf* section 19) does it retain the value a found in the first measurement. This situation is entirely analogous to that in classical physics where an observable, say a position coordinate q, is a function of the time, $q=q(t)$, its variation being determined by the equations of motion.

It follows from Postulate IV that if the second measurement is to have the certain result a, then the wave function ψ', representing the system immediately after the first measurement, must be an eigenfunction of A belonging to the eigenvalue a. If a is a non-degenerate eigenvalue, this determines the wave function ψ' completely. If a is ν-fold degenerate, we only know that ψ' lies in a ν-dimensional subspace, spanned by ν linearly independent eigenfunctions belonging to the eigenvalue a (we shall take these as the normed orthogonal sequence $\varphi_{r+1}, \varphi_{r+2}, \ldots \varphi_{r+\nu}$, corresponding to $a_{r+1}=a_{r+2}= \ldots a_{r+\nu}=a$), *i.e.* ψ' is of the form

$$\psi'=\frac{\sum\limits_{n=r+1}^{r+\nu} c_n\varphi_n}{\sqrt{\left\{\sum\limits_{n=r+1}^{r+\nu}|c_n|^2\right\}}}, \tag{25}$$

but the coefficients c_n are in general unknown. (The sum in the denominator ensures that ψ' is correctly normed: $\|\psi'\|=1$.) However, if the initial wave function ψ is known, we shall postulate that the wave function ψ' after measuring $\mathfrak{A}$ is also known, the coefficients c_n in (25) being given by

$$c_n=(\varphi_n, \psi) \quad (n=r+1, \ldots r+\nu), \tag{26}$$

[and all other $c_n=0$, as implied in (25)]. In other words, the expansion coefficients of $\varphi_{r+1}, \ldots \varphi_{r+\nu}$ are the same in ψ and ψ'. The reasons for this assumption will become apparent in section 19 dealing with simultaneous measurements of several observables. It follows from (25) and (26) that if ψ is an eigenfunction belonging to the eigenvalue a, then $\psi'=\psi$, *i.e.* the wave function is unchanged by the measurement.

We generalize these results somewhat in

> *Postulate V.* If the result of measuring $\mathfrak{A}$ on a system is a value a lying in $a' \leqslant a \leqslant a''$, then, immediately after measurement, the system is represented by a wave function ψ' of the form
>
> $$\psi'=\frac{\sum c_n\varphi_n}{\sqrt{\{\sum|c_n|^2\}}}, \tag{27}$$

the summations being *only* over the values of n for which $a' < a_n < a''$, but the coefficients c_n are not known. If the initial state ψ of the system, before the measurement, is known, then ψ' is also completely known, the coefficients c_n in (27) being the expansion coefficients of φ_n in ψ, *i.e.*

$$c_n = (\varphi_n, \psi). \tag{28}$$

Postulate V states how one incorporates information about a system, obtained as a result of experiments, into the wave function describing that system. In general, such a measurement may lead to a variety of results and the wave function of the system must allow for all these. But once the experiment has been carried out and one particular result obtained, we can 'discard' most of the wave function, corresponding to all the results which were not observed, and retain only that part which on immediate remeasurement would lead to the same result. Thus Postulate V partly answers the question how we determine the state of a system, *i.e.* its state vector. This determination is only complete if the eigenvalue measured is non-degenerate; in general it will be degenerate and further measurements are required to determine the state completely. We shall return to this problem in section 19 dealing with the simultaneous measurement of several observables.

17. Operators with continuous spectra

In section 4 we dealt with the mathematical aspects of operators with continuous spectra of eigenvalues. We now want to see in what way the physical interpretation given for operators with discrete spectra in section 16 has to be modified to deal with the continuous case.

Suppose the operator $\mathfrak{A}$ has a completely continuous spectrum of eigenvalues a_ν and corresponding eigenfunctions $\varphi(\nu, x)$, the parameter ν labelling these eigenfunctions assuming all values in the interval $-\infty < \nu < \infty$. We have written ν as argument rather than as suffix of the eigenfunctions to stress its continuous nature as distinct from the suffix n $(= 1, 2, 3, \ldots)$ which occurred in the discrete case. A general state $\psi(x)$ of the system will no longer be given by an infinite series, equation (16.14), but by an integral

$$\psi(x) = \int_{-\infty}^{\infty} c(\nu)\varphi(\nu, x)\,d\nu. \tag{1}$$

For a system described by (1) we would expect as the result of measuring $\mathfrak{A}$ any one of the values a_ν, which now form a continuum. Accordingly, in the probability Postulate IV, equation (16.15) becomes replaced by

$$P(a_\alpha, a_\beta) = \int_\alpha^\beta |c(\nu)|^2 d\nu. \tag{2}$$

Equation (2) gives the probability of observing a value a, lying in the interval $a_\alpha < a < a_\beta$, when measuring $\mathfrak{A}$ on a system in a state ψ. For the expectation value of $\mathfrak{A}$ we obtain, analogously to equation (16.17)

$$\langle A\rangle = \int_{-\infty}^{\infty} a_\nu |c(\nu)|^2 d\nu. \tag{3}$$

We see that there exists a close similarity between the discrete and continuous cases. To go from the discrete to the continuous case we replace summations over the integers n by integration over the continuous variable ν.

There exists however one principal difference. If A has a discrete spectrum we can measure it with absolute accuracy, as discussed in section 16. If A has a continuous spectrum this is not possible. We see this at once from equation (2). An infinitely sharp measurement implies $\alpha = \beta$, *i.e.* zero probability. So all we can do in the case of a continuous spectrum is to measure $\mathfrak{A}$ with arbitrarily high precision by making $(\beta - \alpha)$ sufficiently small, but we cannot measure it with absolute precision. Apart from this difference, the interpretation is not altered if one goes from operators with discrete spectra to operators with continuous ones.

In section 6 we stated that

$$P(q_1, q_1 + dq_1; \dots ; q_n, q_n + dq_n) = |\psi(q_1, \dots q_n)|^2 dq_1 \dots dq_n \tag{4}$$

is the probability of finding a system specified by the wave function $\psi(q_1, \dots q_n)$ in a state where the coordinate q_i lies in the range $(q_i, q_i + dq_i)$, $i = 1, \dots n$. Hence the expectation value $\langle q_i\rangle$ is

$$\langle q_i\rangle = \int q_i |\psi(q_1, \dots q_n)|^2 dq_1 \dots dq_n, \tag{5}$$

the integration being over the fundamental domains of the coordinate, $q_1, \dots q_n$. Comparing equations (2) and (3) with (4) and (5) respectively, we see that the probability interpretation of $|\psi|^2 dq_1 \dots dq$, is a special case of equation (2).

18. Functions of operators

We now want to consider some of the properties of linear operators which are particularly important in quantum mechanics.

(*a*) *Functions of operators*

For two operators A and B we define the operator $(aA + bB)$ by

$$(aA + bB)\psi = aA\psi + bB\psi \tag{1}$$

for all complex numbers a and b.

The product (AB) is defined as the operation of successively applying first B and then A, *i.e.*

$$(AB)\psi = A(B\psi).$$

In general

$$AB\psi \neq BA\psi.$$

In this case we say that A and B do not commute: the commutator $[A, B]$, defined by

$$[A, B] \equiv AB - BA, \tag{2}$$

does not vanish. This non-commutability of operators is closely related to the uncertainty principle. (*Cf* section 20.)

If $B = A$, we can successively define powers of A, namely $A^2, A^3, \ldots$, and using (1) we can then define polynomial operators

$$P_N(A) = c_0 + c_1 A + \ldots + c_N A^N \tag{3}$$

where $c_0, c_1, \ldots c_N$ are complex numbers. If we let $N \to \infty$, we can define a general function of an operator $f(A)$. We had examples of such functions of operators in equations (16.9) or (16.13).

If φ_n is an eigenfunction of A, corresponding to the eigenvalue a_n, *i.e.* $A\varphi_n = a_n\varphi_n$, then

$$f(A)\varphi_n = f(a_n)\varphi_n, \tag{4}$$

and, generally, two functions $f(A)$ and $g(A)$ always commute

$$[f(A), g(A)] = 0$$

and have a common set of eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots$ and the eigenvalues $f(a_1), f(a_2), \ldots$ and $g(a_1), g(a_2), \ldots$ respectively. All these results follow by considering an individual term in the formal power series expansions of $f(A)$ and $g(A)$:

$$f(A) = c_0 + c_1 A + \ldots + c_N A^N + \ldots, \textit{etc.}$$

Thus (4) follows from

$$A^N\varphi_n = A^{N-1}(A\varphi_n) = A^{N-1}(a_n\varphi_n) = a_n(A^{N-1}\varphi_n) = \ldots = (a_n)^N\varphi_n.$$

(*b*) *Adjoint operators*

In section 16 we defined a Hermitian or self-adjoint operator A by one of the equations (16.4), (16.5) or (16.6). Let us take (16.6) as definition: for all ψ and χ

$$(\chi, A\psi) = (A\chi, \psi). \tag{5}$$

In general an operator will not be Hermitian or self-adjoint. If we can define an operator $A^\star$ such that

$$(\chi, A\psi) = (A^\star\chi, \psi), \tag{6}$$

we call $A^\star$ the *adjoint* operator to A. If A is self-adjoint, then

$$A^\star = A. \tag{7}$$

It follows at once that $(A^\star)^\star = A$, *i.e.* the adjoint of $A^\star$ is A, and furthermore that

$$\left.\begin{aligned} (A+B)^\star &= A^\star + B^\star, \\ (aA)^\star &= \bar{a}A^\star, \end{aligned}\right\} \tag{8}$$

and that

$$(AB)^\star = B^\star A^\star. \tag{9}$$

These relations all follow at once from (6), *e.g.* one gets equation (9):

$$(\chi, AB\psi) = (A^\star\chi, B\psi) = (B^\star A^\star\chi, \psi).$$

If A and B are self-adjoint, then (9) gives

$$(AB)^\star = BA. \tag{10}$$

If in addition A and B commute, $[A, B] = 0$, then (AB) is self-adjoint as well; from (10)

$$(AB)^\star = AB. \tag{11}$$

If $f(x)$ is a complex function[*], then the adjoint of $f(A)$ is

$$[f(A)]^\star = \bar{f}(A^\star) \tag{12}$$

and if f is a real function

$$[f(A)]^\star = f(A^\star). \tag{13}$$

If in addition A is Hermitian, then $f(A)$ is Hermitian too, as one sees at once from equation (13). Equation (12) follows by considering a typical term in the expansion of $f(A)$ and repeatedly using equations (8) and (9), with $B = A$.

(*c*) *Unitary and inverse operators*

So far we have considered Hermitian operators which correspond to the observables. We now want to consider a different type of operator which plays a part analogous to the coordinate transformations of classical physics and is consequently very important.

We define the *unitary* operator U by the conditions

$$UU^\star = U^\star U = \mathbf{I} \tag{14}$$

where $U^\star$ is the adjoint of U and $\mathbf{I}$ is the identity or unit operator, which leaves any vector ψ unchanged,

$$\mathbf{I}\psi = \psi.$$

From (14) it follows that for any vectors ψ and χ

$$(\chi, U^\star U\psi) = (\chi, \psi),$$

and hence

$$(U\chi, U\psi) = (\chi, \psi), \tag{15}$$

i.e. a unitary transformation leaves the inner product of any two vectors *invariant*. This is the most important property of unitary operators and one can use it to define unitary operators. It follows in particular from (15) that

$$\|U\psi\| = \|\psi\| \tag{16}$$

i.e. unitary transformations preserve the length of a vector, and if ψ and χ are orthogonal, then

$$(U\chi, U\psi) = (\chi, \psi) = 0, \tag{17}$$

---

[*] If $f(x)$ is a complex function, it can be written $f(x) = u(x) + iv(x)$, where $u(x)$ and $v(x)$ are real functions. The complex conjugate function is defined as $\bar{f}(x) = u(x) - iv(x)$. If $f(x)$ is a real function, then $\bar{f}(x) = f(x)$.

i.e. the transformed vectors are still orthogonal. We see that unitary transformations have properties similar to rotations in ordinary space. They are in fact the corresponding generalizations to function space.

If A is Hermitian and f a real function, then $\boldsymbol{U} = \exp[if(A)]$ is a unitary operator. For from (12) it follows that $U^* = \exp[-if(A)]$, and hence equations (14) and (15) are satisfied identically. [Equation (15) becomes

$$(\exp[if(A)]\chi, \exp[if(A)]\psi) = (\exp[-if(A)] \exp[if(A)]\chi, \psi) = (\chi, \psi).]$$

We see now that the operators S_r and T_r introduced in (16.9) and (16.13) are unitary operators and equation (16.10) is only a special case of (15).

From equation (14) it follows that if $\psi = U\varphi$, then $\varphi = U^*\psi$, *i.e.* for a unitary operator U the adjoint U^* has the property of being in a sense an inverse operator to U. This suggests that for any operator A we should try to define an *inverse* operator which we shall denote by A^{-1} and which has the property that

$$AA^{-1} = A^{-1}A = \mathbf{I} \tag{18}$$

so that $\psi = A\varphi$ implies $\varphi = A^{-1}\psi$. From (15) and (18) it follows that any unitary operator U possesses an inverse U^{-1} and that

$$U^{-1} = U^*. \tag{19}$$

19. Simultaneous measurability of observables

So far we have dealt with the measurement of only one observable A. This does not in general specify a system completely and we must now consider the measurement of several observables.

According to classical mechanics, we can determine the state of a system with n degrees of freedom at time t by measuring the $2n$ position and momentum coordinates $q_i(t)$, $p_i(t)$, $i = 1, \ldots n$. We can think of these measurements as either simultaneous or as carried out in very rapid succession so that the coordinates have not varied appreciably, on account of their temporal development according to the equations of motion, during the interval of time required for all the measurements. But it is inherent in the concepts of classical physics that the measurements do not affect the values of the variables describing the system. Thus, the order of measuring the $2n$ coordinates is immaterial and remeasuring any one of them sufficiently quickly necessarily leads to the same result. Furthermore, all conceivable observables are functions of the $2n$ generalized coordinates; from these, any other quantity can be calculated, and specifying the system by $2n$ different generalized coordinates, to begin with, does not lead to any new or essentially different results.

In quantum mechanics the situation is entirely different. It is no longer possible to specify the simultaneous values of all observables of a system. In general, observables are mutually incompatible, measuring one of them partly or completely destroys any knowledge about the others. If two observables A and B possess a mutual eigenfunction φ, *i.e.* $A\varphi=a\varphi$ and $B\varphi=b\varphi$, then, in the state φ, A has the value a and B the value b, and it follows from Postulates IV and V (section 16) that successive measurements (performed sufficiently quickly for the time variation of the system to be unimportant) would give these results. However, if A and B have no common eigenfunctions, successive alternate measurements of A and B, which necessarily leave the system in eigenstates of A and B respectively, destroy the information obtained earlier. We shall deal with such incompatible observables in section 20. They are closely related to the uncertainty principle. In this section we consider under what conditions one can ascribe simultaneous values to several observables.

To begin with, we again consider two operators A and B with completely discrete spectra of eigenvalues. Let $a_1<a_2<\dots<a_r<\dots$ and $b_1<b_2<\dots b_s<\dots$ be the *distinct* eigenvalues of A and B respectively, which may of course be degenerate. We define the compatibility or simultaneous measurability of A and B as follows:

> *Definition 1.* If A and B are measured on a system, in quick succession, with the results a_r and b_s, and an immediate remeasurement of A or B necessarily reproduces these results, *whatever the initial state of the system*, then we say that A and B are *compatible* or *simultaneously measurable*.

From the postulates of section 16 the following theorem can be deduced which is of fundamental importance.

Theorem 1. A necessary and sufficient condition for A and B to be compatible is for them to possess a complete normed orthogonal set of simultaneous eigenfunctions.

(*a*) The condition is necessary, *i.e.* if A and B are compatible, there exists such a set.

Suppose the system is initially in a state ψ. After a measurement of A, with result a_r, it is in an eigenstate ξ_r of A belonging to the eigenvalue a_r. Expanding ξ_r in terms of the eigenfunctions of B, we have

$$\xi_r=\sum_{s=1}^{\infty}c_s\eta_s, \tag{1}$$

where η_s is a normed eigenfunction of B belonging to the eigenvalue b_s. (In general, b_s will be degenerate so that there exist several linearly independent eigenfunctions belonging to this eigenvalue b_s. But the particular linear combination $c_s\eta_s$ occurring in (1) is of course uniquely determined by ξ_r.) A measurement of B on the system in state (1) (*i.e.* we measure B immediately after we have measured A) with the result b_s leaves the system in the state η_s. But since A and B are compatible, a second measurement of A must necessarily give a_r again, *i.e.* η_s is a simultaneous eigenstate of A and B. Hence any eigenstate ξ_r of A can be expanded in a series of mutual eigenfunctions of A and B. Since any state ψ of the system can be written as a superposition of eigenstates of A, it can also be expanded in terms of simultaneous eigenfunctions of A and B, *i.e.* the set of mutual eigenfunctions is complete.

(*b*) The condition is sufficient, *i.e.* if there exists a complete set of mutual orthonormal eigenfunctions $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ then A and B are compatible.

Since the set $\tau_1, \tau_2, \ldots, \tau_n, \ldots$ is complete, any state ψ of the system can be expanded in terms of this set,

$$\psi=\sum_{n=1}^{\infty}c_n\tau_n \qquad \left(\sum_{n=1}^{\infty}|c_n|^2=1\right).$$

After measuring A, with result a_r, the state of the system is

$$\frac{\sum c_n\tau_n}{\sqrt{\{\sum|c_n|^2\}}},$$

the summations being only over those values of n for which τ_n is an eigenfunction of A belonging to the eigenvalue a_r:

$$A\tau_n=a_r\tau_n.$$

A subsequent immediate measurement of B, result b_s, leads to the state

$$\psi'=\frac{\sum c_n\tau_n}{\sqrt{\{\sum|c_n|^2\}}},$$

the summations now being *only over those values of* n *for which* τ_n *is a mutual eigenfunction of* A *and* B *belonging to the eigenvalues* a_r *and* b_s:

$$A\tau_n=a_r\tau_n, \quad B\tau_n=b_s\tau_n.$$

Hence remeasuring A gives a_r, and A and B are compatible.

Theorem 1 gives the generalization of Postulate IV (section 16) for several observables. If $\psi=\sum c_n\tau_n$, then the probability of observing the values a_r and b_s simultaneously for A and B is $\sum|c_n|^2$, the summation again being only over those values of n for which $A\tau_n=a_r\tau_n$ and $B\tau_n=b_s\tau_n$.

Theorem 1 has a simple geometrical interpretation. The complete normed orthogonal set $\tau_1, \tau_2, \ldots \tau_n, \ldots$ spans the Hilbert space $\mathfrak{H}$. We can subdivide the whole Hilbert space into mutually orthogonal subspaces $\mathfrak{M}\,(a_r, b_s)$ in each of which A and B have definite eigenvalues a_r and b_s for all vectors belonging to them. Each of these subspaces is spanned by the mutual eigenvectors τ_n belonging to the eigenvalues a_r and b_s, and their number specifies the dimensionality of the subspace.

We can write this decomposition as

$$\mathfrak{H}=\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\mathfrak{M}(a_r, b_s). \tag{2}$$

If we know the results a_r and b_s of a measurement of A and B, and the corresponding manifold $\mathfrak{M}(a_r, b_s)$ is many-dimensional, the state ψ' after the measurement depends on the coefficients c_n of the relevant τ_n in the original wave function ψ. On the other hand, if $\mathfrak{M}(a_r, b_s)$ is one-dimensional, $\psi'=\tau_n$, *i.e.* the state after measurement is independent of the coefficients c_n of the original wave function ψ. (It is of course implied that $c_n \neq 0$, so that we can obtain the results a_r and b_s postulated.)

In section 18 we saw that the two operators A and B do not, in general, commute. Correspondingly, we have:

Definition 2. Two operators A and B *commute*, if for all wave functions ψ

$$AB\psi=BA\psi. \tag{3}$$

If this equation is to be identically satisfied for all ψ, it follows that the commutator vanishes identically,

$$[A, B]=0, \tag{4}$$

i.e. it transforms any wave function into zero. Two operators A and B commute if (3) holds for every member of a complete set of functions $\tau_1, \tau_2, \ldots \tau_n, \ldots$ For any other function ψ can be written as a linear superposition of these functions $\tau_1, \tau_2, \ldots \tau_n, \ldots$. We can then prove the very important, purely mathematical

Theorem 2. A necessary and sufficient condition for the operators A and B to commute,

$$[A, B]=0, \tag{4}$$

is the existence of a complete normalized orthogonal set of *mutual* eigenfunctions, $\tau_1, \tau_2, \ldots \tau_n, \ldots$.

(*a*) The condition is necessary, *i.e.* (4) implies the existence of the set $\tau_1, \tau_2, \ldots \tau_n, \ldots$.

Take any eigenfunction ξ_r of A belonging to the eigenvalue a_r. As in equation (1), it can be written

$$\xi_r=\sum_{s=1}^{\infty}c_s\eta_s, \tag{1}$$

where η_s is a normed eigenfunction of B belonging to the eigenvalue b_s. From (4) and section 18, it follows that for any function f, $Af(B)=f(B)A$, *i.e.*

$$Af(B)\xi_r=f(B)A\xi_r=f(B)a_r\xi_r,$$

hence

$$(A-a_r)f(B)\xi_r=0.$$

From equations (1) and (18.4) it follows that

$$\sum_{s=1}^{\infty}(A-a_r)f(b_s)c_s\eta_s=0,$$

and since f is an arbitrary function, it follows that

$$(A-a_r)\eta_s=0,$$

i.e. all eigenfunctions η_s occurring in (1) are also eigenfunctions of A belonging to the eigenvalue a_r. The proof then follows exactly as in part (a) of Theorem 1.

(b) The condition is sufficient, *i.e.* the existence of a complete set $\tau_1, \tau_2, \ldots \tau_n \ldots$ of mutual eigenfunctions implies (4).

If τ_n belongs to the eigenvalues a_n and b_n of A and B, then

$$AB\tau_n=b_nA\tau_n=a_nb_n\tau_n,$$
$$BA\tau_n=a_nB\tau_n=a_nb_n\tau_n,$$

and the commutability of A and B follows at once.

We can now combine theorems 1 and 2 into the general

Theorem 3. Of the three conditions:

(α) A and B are compatible,
(β) A and B possess a complete normed orthogonal sequence of simultaneous eigenfunctions,
(γ) A and B commute,
any one implies the other two.

Suppose T is a Hermitian operator with the complete normed orthogonal set of eigenfunctions $\tau_1, \tau_2, \ldots \tau_n \ldots$ and corresponding eigenvalues $t_1 < t_2 < \ldots < t_n < \ldots$. From section 18 it follows that if we define two operators A and B which are functions of T

$$A=f(T), \quad B=g(T),$$

then A and B have the same complete set of eigenfunctions $\tau_1, \tau_2, \ldots \tau_n, \ldots$ and the eigenvalues $f(t_1), f(t_2), \ldots$, and $g(t_1), g(t_2), \ldots$, respectively. It can conversely be proved that if two operators A and B commute, they are functions of a third operator T, $A=f(T)$ and $B=g(T)$. We shall not prove this converse. But we state it, as it represents a particularly simple and lucid formulation of the above results.

The above ideas can easily be generalized, firstly to any number of observables with discrete spectra, secondly to operators with continuous spectra. In the latter case the only change that occurs is that we no longer have measurements of absolute accuracy, but only of arbitrarily high accuracy, the situation being essentially the same as in section 17.

Consider, for example, the position and momentum coordinates (x_1, x_2, x_3) and (p_1, p_2, p_3) of a particle. Since $p_r = -i\hbar\partial/\partial x_r$, we have at once that (*cf* Exercise IV.8)

$$[x_r, x_s] = 0, \quad [p_r, p_s] = 0, \quad (r, s = 1, 2, 3), \tag{5a}$$

but

$$[x_r, p_s] = i\hbar\delta_{rs}, \quad (r, s = 1, 2, 3). \tag{5b}$$

These equations are of course always to be understood as operator equations: the two sides have the same effect when applied to a wave function. In contrast to equations (5) the position and momentum coordinates of two different particles always commute. If (X_1, X_2, X_3), (P_1, P_2, P_3) are the position and momentum coordinates of a second particle, then

$$[x_r, X_s] = [p_r, P_s] = [x_r, P_s] = [p_r, X_s] = 0.$$

Consider the complete set of eigensolutions of a particle moving in a central field of force, equation (13.27).

$$\psi_{E_{n,l}, l, m}(r, \theta, \varphi) = R_{n,l}(r) Y_l^m(\theta, \varphi) \tag{6}$$

is a simultaneous eigenfunction of l_z, $\mathbf{l}^2$ and the Hamiltonian operator H, the eigenvalues being $\hbar m$, $\hbar^2 l(l+1)$ (equations (14.5) and (14.6)) and $E_{n,l}$, determined from (12.7). It follows that these three operators commute,

$$[H, l_z] = 0 \quad [H, \mathbf{l}^2] = 0 \tag{7}$$

and

$$[\mathbf{l}^2, l_z] = 0 \tag{8}$$

which can easily be proved directly (*cf* Exercise IV.9).

We can now say how the state of a system at a given instant of time is determined. As we have seen we cannot increase our knowledge of the state of a system by successive measurements of any arbitrary observables. Rather, having measured A, we may only measure a second observable B if it commutes with A, and so on. After measurement of A, with the $\mathcal{N}(a_r)$-fold degenerate eigenvalue a_r as result, the wave function of the system necessarily lies in the $\mathcal{N}(a_r)$-dimensional subspace $\mathfrak{M}(a_r)$ of Hilbert space, comprising all eigenfunctions of A belonging to a_r. After a subsequent immediate measurement of an observable B which commutes with A, with result b_s, the wave function of the system lies in a subspace $\mathfrak{M}(a_r, b_s)$ which is a subspace contained in $\mathfrak{M}(a_r)$, whose dimensionality $\mathcal{N}(a_r, b_s)$ is necessarily not greater than $\mathcal{N}(a_r)$:

$$\mathcal{N}(a_r, b_s) \leqslant \mathcal{N}(a_r).$$

It comprises all simultaneous eigenfunctions of A and B belonging to a_r and b_s. In this way we continue measuring more and more commuting observables $A, B, C, \ldots$. At each step we reduce the dimensionality of the manifold within which the wave function may lie; the degeneracy of the state is reduced; the number of possible ways of resolving the state by a further measurement is reduced. Finally we have a set of commuting operators $A, B, \ldots Q$, such that to any set of eigenvalues $a_r, b_s, \ldots q_z$, there exists *exactly one* eigenfunction, $N(a_r, b_s, \ldots q_z)=1$. This state cannot be resolved further by more measurements. We call such an observation *maximal* or *complete*, and the corresponding set of observables (or operators) a complete set of commuting observables (or operators). To determine the state of a system we must carry out a maximal observation. Thus, in the above example of a particle in a central field, H, $\mathbf{l}^2$ and l_z constitute a complete set of commuting observables, but not H and $\mathbf{l}^2$ by themselves, for example. For after a measurement of H and $\mathbf{l}^2$, with the results E_n, and $\hbar^2 l(l+1)$, there is still a $(2l+1)$-fold degeneracy of the wave function corresponding to the $(2l+1)$ possible values of l_z.

From our derivation of a complete set of observables $A, B, \ldots Q$, it is clear that after a maximal observation we do not know the exact values of all possible observables. We cannot ascribe a definite value to any observable W which does not commute with all observables $A, B, \ldots Q$ of the set. Only if W commutes with $A, B, \ldots Q$ can we ascribe a definite value to it. Any function $W=f(A, B, \ldots Q)$ is an example of such an observable. But we do of course know the exact wave function of the system after a maximal observation, hence we can calculate the statistics (mean values, probability distributions *etc*) of *any* observable. It is furthermore clear that there is not one unique complete set of commuting observables. Thus, if we had taken an operator W for the first measurement, which does not commute with every operator of the set $A, B, \ldots Q$, we would have obtained a different complete set of commuting observables. For the particle in a central field, H, $\mathbf{l}^2$ and l_z form a complete set of observables. Since l_x and l_y do not commute with l_z, we cannot determine their values simultaneously with l_z. But H, $\mathbf{l}^2$, l_x or H, $\mathbf{l}^2$, l_y would be equally acceptable complete sets of commuting operators.

We consider now the operators which commute with the Hamiltonian function of a system. We shall restrict ourselves to the case, sufficiently general for most problems, that the Hamiltonian operator H does not contain the time explicitly ($\partial H/\partial t=0$). In this case the explicit solution of the Schrödinger equation, giving the temporal development of the wave function, is [equations (16.12) and (16.13)]

$$\psi_t = T_t\psi_0, \tag{9}$$

where $T_t=\exp(-iHt/\hbar)$ is a unitary operator. Suppose A is an operator which does not contain the time explicitly, and which commutes with the Hamiltonian, $[A, H]=0$. It follows from section 18 that A then also commutes with T_t which is a function of H, $[A, T_t]=0$, and since T_t is unitary so that $T_t^{-1}=T_t^{\star}$, we have that

$$A=T_t^{\star}AT_t. \tag{10}$$

If ψ_0 represents the state of the system at time $t=0$, the expectation value of A at $t=0$ is

$$\langle A\rangle_0=(\psi_0, A\psi_0), \tag{11}$$

its expectation value at time t is

$$\langle A\rangle_t=(\psi_t, A\psi_t), \tag{12}$$

and from equations (10), (11) and (12) it follows at once that

$$\langle A\rangle_t=(T_t\psi_0, AT_t\psi_0)=(\psi_0, T_t^{\star}AT_t\psi_0)=(\psi_0, A\psi_0)=\langle A\rangle_0, \tag{13}$$

i.e. the expectation value of A remains constant throughout time, while the system is undisturbed and its temporal development is described by (9). In exactly the same way one proves, that if

$$A\varphi_0=a\varphi_0 \tag{14}$$

and $\varphi_t=T_t\varphi_0$, then

$$A\varphi_t=a\varphi_t, \tag{15}$$

i.e. if the system is in an eigenstate of A belonging to the eigenvalue a at time $t=0$, then, under the undisturbed temporal development, it remains in an eigenstate of A belonging to the same eigenvalue a. Thus A has a constant expectation value for this system, and if the system is in an eigenstate of A, then it remains in that eigenstate. Hence we call an operator A which does not contain the time explicitly and which commutes with the Hamiltonian a *constant of the motion*, and the above results are the quantum mechanical analogues of the classical conservation theorems. An important consequence of the above results is that, if we specify the system by a complete set of commuting operators containing the Hamiltonian H as one of its members, a complete observation gives us the values of these observables not only immediately after the measurement, but for all subsequent times, as long as the system is not disturbed.

In the case of a particle in a central field of force, we see now that the magnitude of the angular momentum $|\mathbf{l}|$ and its z-component l_z are constants of the motion [equations (7)] as already mentioned in section 14, analogously to the classical conservation of angular momentum in this case. But unlike the classical case, we cannot say anything about l_x and l_y. For though relations analogous to (7) hold,

$$[H, l_x]=[H, l_y]=0,$$

as follows simply from symmetry arguments (the z-direction is in no way preferred), we cannot specify l_x and l_y simultaneously with l_z, since these operators do not commute with l_z, and hence there is no sense in talking of their constancy.

20. Uncertainty principle

We now come to consider two operators A and B which do not commute. We know already that the corresponding observables are not compatible. However, it may happen that A and B have some eigenfunctions in common and for these states one can ascribe definite values to A and B, but they may not possess a complete set of simultaneous eigenfunctions. If ψ is a mutual eigenstate, then $A\psi = a\psi$, and $B\psi = b\psi$, *i.e.* $(AB - BA)\psi = 0$, *i.e.* ψ is an eigenfunction of the operator $[A, B]$, belonging to the eigenvalue 0. This situation is illustrated by the angular momentum operators l_x, l_y, l_z. These do not commute with each other, but the state ψ_0^0, equation (14.4), is a simultaneous eigenfunction of all three operators with the common eigenvalue 0, and $[l_x, l_y]\psi_0^0 = 0$, *etc.*

If A and B are two self-adjoint operators which do not commute, and ψ is a normed wave function of the system for which

$$[A, B]\psi = ia\psi, \tag{1}$$

then the mean square deviations [equation (16.19)] ΔA and ΔB satisfy

$$\Delta A\,\Delta B \geq \tfrac{1}{2}|a|. \tag{2}$$

The operator $[A, B]$ is not Hermitian, since

$$[A, B]^* = (AB - BA)^* = BA - AB = -[A, B].$$

Taking the adjoint operators of (1), we get

$$-[A, B] = -i\bar{a}$$

i.e. $\bar{a} = a$; a is real. [This was of course our object in writing ia rather than a in (1).]

To prove (2), we write

$$\alpha = A - \langle A\rangle, \quad \beta = B - \langle B\rangle$$

where $\langle A\rangle = (\psi, A\psi)$, *etc.* One then gets from (1)

$$\alpha\beta - \beta\alpha = ia \tag{3}$$

since $\langle A\rangle$ and $\langle B\rangle$ are ordinary numbers which commute with all operators. Consider the imaginary part of the inner product $(\beta\psi, \alpha\psi)$.

$$\begin{aligned} 2\,Im(\beta\psi, \alpha\psi) &= (\beta\psi, \alpha\psi) - (\alpha\psi, \beta\psi) \\ &= (\psi, [\beta\alpha - \alpha\beta]\psi) \\ &= -ia, \end{aligned} \tag{4}$$

using equation (3). Taking moduli of (4), we have at once

$$\begin{aligned} |a| &= 2|Im(\beta\psi, \alpha\psi)| \\ &\leq 2|(\beta\psi, \alpha\psi)| \\ &\leq 2\|\beta\psi\|\,\|\alpha\psi\| \\ &= 2\Delta A\,\Delta B, \end{aligned}$$

where the last step but one follows from Schwartz's inequality (*cf* Exercise I.7).

Equation (2) is a statement of the *uncertainty principle* first derived by Heisenberg. It expresses the fundamental limitation to the accuracy with which one may hope to know the simultaneous values of two incompatible observables. If one calculates the product of the 'uncertainties' $\Delta A \, \Delta B$, for a given state ψ, one finds that in general the inequality holds. Only under particularly favourable conditions does the equality hold.

From equations (19.5b) we obtain an uncertainty relation between the position and momentum coordinates of a particle,

$$\Delta x_r \, \Delta p_r \geqslant \frac{\hbar}{2}, \quad (r=1, 2, 3). \tag{5}$$

Introducing the velocity vector v_r, we can write (5)

$$\Delta x_r \, \Delta v_r \geqslant \frac{\hbar}{2m}, \tag{6}$$

where m is the mass of the particle considered. Equation (6) shows why the uncertainty principle is not of importance for macroscopic bodies. Since $\hbar \simeq 10^{-27}$ erg sec, $\hbar/2m$ is exceedingly small in this case.

That we cannot know the position and momentum of a particle accurately at the same time, implies that any accurate position measurement partly alters the momentum of the particle in a manner which cannot be completely specified. One can theoretically analyse experiments to confirm this and to verify (5), qualitatively at least. Heisenberg[16] has given a rigorous detailed quantum-mechanical analysis of such experiments. Here we shall limit ourselves to a simple example.

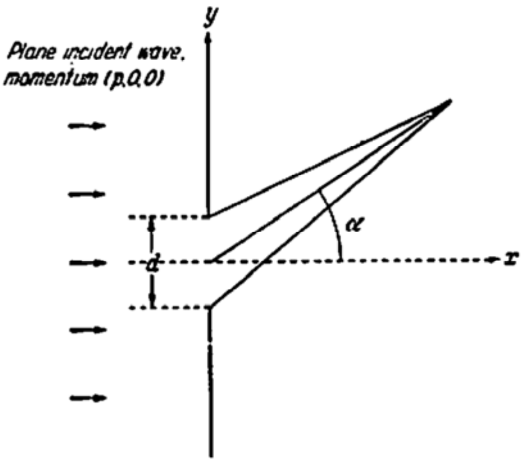

Figure 7. Position measurement of a plane wave of particles

We consider the position measurement of a particle whose momentum $(p, 0, 0)$ is known accurately. Such a particle is described by a plane wave, *i.e.* its position is completely unknown. To determine the position of the particle in the plane $x=0$, we place a partition with a slit of width d in this plane, as indicated in *Figure 7*. After passing through the slit the y-coordinate of the particle is known with an uncertainty equal to the width of the slit,

$$\Delta y \simeq d. \tag{7}$$

But in passing the slit the beam suffers diffraction, and hence its momentum, after going through the slit, is no longer $(p, 0, 0)$, but has a component in the y-direction. The first minimum of the diffraction pattern occurs at an angle α such that

$$d \sin \alpha = \lambda. \tag{8}$$

Corresponding to this deflection, the particle has a y-component of momentum

$$p_y = p \sin \alpha = \frac{h}{\lambda} \sin \alpha,$$

and hence there is an uncertainty in the y-component of momentum (corresponding to all angles of deflection between $-\alpha$ and $+\alpha$)

$$\Delta p_y \simeq \frac{2h}{\lambda} \sin \alpha. \tag{9}$$

Combining equations (7), (8) and (9), we obtain

$$\Delta y \Delta p_y \simeq 2h,$$

as verification of the uncertainty relation. It is clear that this sort of analysis is rather qualitative. For example, we could equally well have taken an angle α defined by

$$d \sin \alpha = \lambda/2,$$

corresponding to interference between the extreme rays at the two edges, to define the uncertainty in the momentum. We would then have obtained

$$\Delta y \Delta p_y \simeq h.$$

We obtain another example of the uncertainty relations by considering the orbital angular momentum of a particle. Using the commutation relations (19.5), one easily shows that the angular momentum operators l_x, l_y and l_z satisfy the commutation relations

$$[l_x, l_y] = i\hbar l_z \tag{10}$$

and two similar relations obtained by permuting x, y and z cyclically (*cf* Exercise IV.13). If we consider the state ψ_l^m, equation (14.4), (10) gives

$$[l_x, l_y]\psi_l^m = i\hbar^2 m \psi_l^m,$$

and hence, from (2), the uncertainty relation

$$\Delta l_x \Delta l_y \geqslant \frac{\hbar^2 |m|}{2} \tag{11}$$

follows.

Another way of looking at this uncertainty relation is to say that in the state ψ_l^m we do not know l_x or l_y precisely, but only that

$$l_x^2 + l_y^2 = \mathbf{l}^2 - l_z^2$$

has the exact value $\hbar^2(l^2 + l - m^2)$. Thus, we know the angle α, equation (16.24), between the vector $\mathbf{l}$ and the z-axis but not the direction of the component $\sqrt{(l_x^2 + l_y^2)}$ in the (x, y)-plane. We illustrate this by the vector diagram in *Figure 8* showing the angular momentum vector in the case $l = 5$. This figure is axially symmetric about the z-axis and must be imagined rotated about the z-axis as indicated by the dotted circles. The azimuthal angle φ of the $(\mathbf{l}, z)$-plane with the (x, z)-plane is not known.

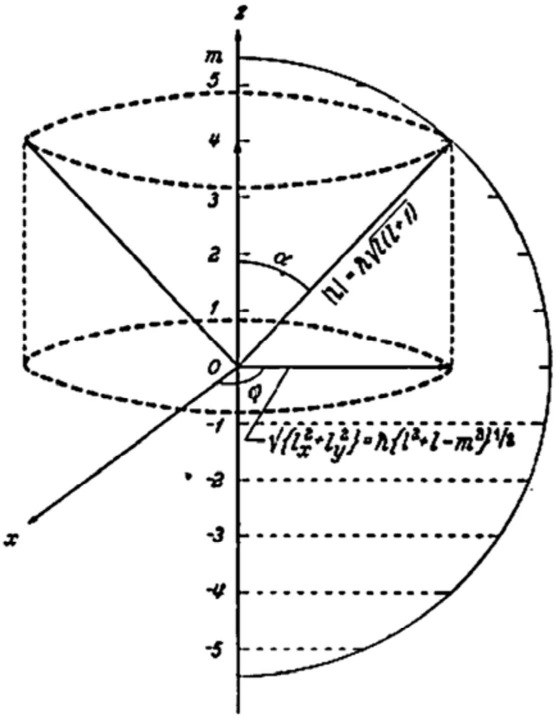

Figure 8. Vector diagram of orbital angular momentum of a single particle for the case $l = 5$

Since from (16.19),

$$\langle l_x^2 \rangle = \langle l_x \rangle^2 + (\Delta l_x)^2,$$

and from symmetry considerations

$$\Delta l_x = \Delta l_y,$$

it follows from (11) that the value of $(l_x^2 + l_y^2)$ in the state ψ_l^m is

$$\hbar^2(l^2 + l - m^2) \geqslant (\Delta l_x)^2 + (\Delta l_y)^2 \geqslant \hbar^2 |m|.$$

Hence if $m = l$, *i.e.* in the state, ψ_l^l, $(l_x^2 + l_y^2)$ has the minimum value compatible with the uncertainty principle. l_z has correspondingly its maximum value $\hbar l$, and on account of the uncertainty principle it cannot attain the value $\hbar\sqrt{l(l+1)}$ as discussed in connection with equation (16.24).

An uncertainty relation, similar to (5), exists between energy and time

$$\Delta t \, \Delta E \simeq \hbar. \tag{12}$$

This relation cannot be derived rigorously within the framework of non-relativistic quantum mechanics which we have developed. We have treated the time on an entirely different footing from the other coordinates. Whereas these latter were described by operators operating on wave functions, the time was always a parameter and as such an ordinary number. Hence it commutes with all operators. But equation (2), of which (12) is to be a special case, was derived for non-commuting operators.

That an uncertainty relation (12) is to be expected can be seen formally by going over to a relativistic notation. In the special theory of relativity one supplements the position vector (x_1, x_2, x_3) of a particle by the 'time coordinate' $x_4 = ict$, to form the 'four-vector' (x_1, x_2, x_3, x_4), and similarly one forms the four-vector (p_1, p_2, p_3, p_4) by adding to the momentum coordinates (p_1, p_2, p_3) the energy component $p_4 = iE/c$. Equation (12) is then the generalization of (5) to include $r = 4$ in the relativistic case. Furthermore with this notation we can write equation (6.3) in the form

$$p_4 = -i\hbar \frac{\partial}{\partial x_4},$$

similar to equation (6.1a),

$$p_r = -i\hbar \frac{\partial}{\partial x_r}, \quad (r = 1, 2, 3),$$

showing that x_4 and p_4 satisfy the same commutation relations as x_r and p_r $(r = 1, 2, 3)$, *i.e.* we replace equations (19.5) by

$$\left.\begin{array}{l} [x_r, x_s] = 0 \\ [p_r, p_s] = 0 \\ [x_r, p_s] = i\hbar\delta_{rs} \end{array}\right\} \quad (r, s = 1, \ldots 4).$$

We must now interpret equation (12). It implies a limitation on the accuracy with which the energy of a system can be measured in a finite time. For an absolutely accurate determination of the energy $(\Delta E = 0)$ we must take infinitely long over the measurement. If the state we measure is stable, we can indeed determine its energy level very accurately, as we can take as long as we like over the measurement. But suppose now the state is unstable with a life time τ (*i.e.* the transition probability per unit time is $\gamma = 1/\tau$). In this case the average time available for an energy measurement of the state is τ; consequently there is an uncertainty ΔE in the energy determination, given by (12), of

$$\Delta E \simeq \hbar/\tau = \hbar\gamma. \tag{13}$$

If we introduce the circular frequency ω, corresponding to the energy $E=\hbar\omega$, then (13) implies an uncertainty in the frequency

$$\Delta\omega \simeq \gamma. \tag{14}$$

In practice, energy levels of systems are often observed spectroscopically. Equation (14) then tells us that for a finite life time τ, the spectral line has a finite width on account of the uncertainty principle.

We have already met spectral lines of finite width in connection with wave trains of finite length (Exercise I.6 and Chapter II), and we use the same method of analysis in the present case. This approach is particularly useful in certain problems of nuclear physics where one considers unstable excited nuclear states. (This analysis leads to the Breit–Wigner formula.) The wave function

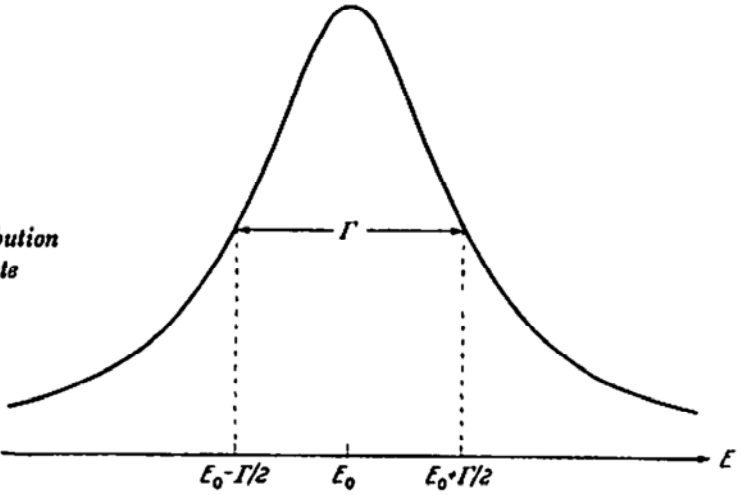

Figure 9. Energy distribution of a non-stationary state

$$\psi_t=\psi_0 e^{-iE_0t/\hbar}$$

represents a stationary state with energy E_0, its amplitude is constant in time, $|\psi_t|^2=|\psi_0|^2$. If we consider a time-dependent wave function

$$\psi_t=\psi_0 e^{-\gamma t/2} e^{-iE_0t/\hbar}, \tag{15}$$

this represents a 'decaying' state, its intensity decreases exponentially,

$$|\psi_t|^2=|\psi_0|^2 e^{-\gamma t},$$

and $\tau=1/\gamma$ is a time characteristic for the rate of decay. Equation (15) does not represent one stationary state, but we can write it as a Fourier integral, *i.e.* as a superposition of stationary states; *i.e.* we write

$$\psi_t=\int_{-\infty}^{\infty} A(E) e^{-iEt/\hbar} dE, \tag{16}$$

where $A(E)$ is the amplitude of the stationary state with energy E in (15). $A(E)$ is given by Fourier's integral theorem, and the corresponding intensity distribution is

$$|A(E)|^2=\frac{|\psi_0|^2}{4\pi^2}\frac{1}{(E-E_0)^2+\Gamma^2/4} \tag{17}$$

where $\Gamma = \hbar\gamma$. This energy spectrum is shown in *Figure 9*. Γ is the half-width of the level, *i.e.* the maximum intensity occurs at E_0, and this intensity has fallen off to half the maximum value at energies $E_0 \pm \Gamma/2$. We see that it is reasonable to take the energy uncertainty of the state (15) as $\Delta E \simeq \Gamma = \hbar\gamma$, in agreement with equation (13).

EXERCISES IV

IV.1. The Hamiltonian function of a system is

$$H = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2) + V(q_1, q_2, q_3); \qquad \text{(IV.1)}$$

obtain Hamilton's equations and interpret them.

IV.2. The motion of a point mass m in a central field $V(r)$ is described in spherical polar coordinates, equations (12.2). If these are taken as generalized coordinates, define the corresponding generalized momenta, the Hamiltonian function, and derive Hamilton's equations.

IV.3. In section 16 we showed that the eigenvalues of a Hermitian operator are all real. Prove the converse, *i.e.* that an operator with a complete set of real eigenvalues is Hermitian.

IV.4. Prove equation (16.22).

IV.5. Prove that, if U and V are unitary operators, then (*i*) U^*, (*ii*) (UV) and (*iii*) (UV^{-1}) are unitary operators.

IV.6. If the operators A and B have the inverse operators A^{-1} and B^{-1}, show that A^{-1} is unique, and find the inverse of (AB).

IV.7. If $\varphi_1, \varphi_2, \ldots \varphi_n \ldots$ is a complete normalized orthogonal sequence of functions, show that the set

$$\psi_i = U\varphi_n \quad (i = 1, 2, \ldots n, \ldots)$$

generated by the unitary operator U, is also a complete normalized orthogonal set.

IV.8. Prove equations (19.5).

IV.9. Prove equation (19.8) directly.

IV.10. If $\psi(x)$ represents a non-degenerate energy eigenfunction of a particle, moving in a potential field $V(x)$ of definite parity, prove that the expectation value of the position coordinate x is zero.

IV.11. If a beam of particles is represented by the normed wave function $\psi(\mathbf{r})$, what is the momentum distribution of this beam, *i.e.* what is the probability of a particle having a momentum in the range $\mathbf{p}$ to $\mathbf{p} + d\mathbf{p}$?

IV.12. If A, B and C are any operators, prove the following identities

$$[A, B + C] = [A, B] + [A, C] \qquad \text{(IV.12a)}$$

$$[A, BC] = [A, B]C + B[A, C] \qquad \text{(IV.12b)}$$

$$[BC, A] = B[C, A] + [B, A]C. \qquad \text{(IV.12c)}$$

IV.13. Prove the following commutation relations for the orbital angular momentum operators l_x, l_y, l_z:

$$[l_x, x] = 0, \quad [l_x, y] = i\hbar z, \quad [l_x, z] = -i\hbar y, \tag{IV.13a}$$

$$[l_x, p_x] = 0, \quad [l_x, p_y] = i\hbar p_z, \quad [l_x, p_z] = -i\hbar p_y, \tag{IV.13b}$$

$$[l_x, l_y] = i\hbar l_z, \tag{IV.13c}$$

$$[l_x, \mathbf{l}^2] = [l_x, \mathbf{r}^2] = [l_x, \mathbf{p}^2] = 0. \tag{IV.13d}$$

All the other commutation relations between the components of $\mathbf{r}$, $\mathbf{p}$ and $\mathbf{l}$ can be derived from (IV.13a–d) by permuting the indices x, y and z.

V

MATRIX MECHANICS

In Chapter I we discussed two methods of describing vector spaces or function spaces. The simpler and more direct method is to use the physically significant quantities, such as the vectors, inner products, *etc* directly. Alternatively we can introduce a coordinate system to which all quantities are referred. This latter method is often convenient in practice though it involves a largely arbitrary coordinate system. In this chapter, we shall deduce this form of quantum mechanics using coordinate systems and components from the coordinate-free formulation which we have so far developed. In this way, matrix mechanics, as this formulation of quantum mechanics is often called, will appear as a natural and simple consequence of the previous theory. Historically, it was actually the first form of quantum mechanics, having been developed in 1925 by Heisenberg, Born and Jordan, and independently by Dirac.

21. Matrix representation of wave functions and operators

To obtain matrix mechanics, we choose a complete orthonormal set of functions $\xi_1, \xi_2, \ldots \xi_n, \ldots$ to form a coordinate system or *basis*, spanning our Hilbert space, and expand all wave functions in terms of this basic set of functions. Equations between wave functions are in this way replaced by equations between the expansion coefficients (*i.e.* the components) of wave functions. These equations assume a particularly simple form if we introduce the notation of matrix algebra. We shall, in fact, in this way deduce many of the laws of matrix algebra from previous results for linear operators and wave functions; but it should be stressed that matrix algebra forms a complete and systematic subject, and all its results can be derived using matrix methods only.

For the time being, we shall restrict ourselves to the one coordinate system $\xi_1, \xi_2, \ldots \xi_n, \ldots$ Later we shall also admit other bases, and we shall have to consider the transformations between these equally acceptable coordinate systems. We shall always limit ourselves to orthonormal sets as these are most convenient in most applications.

We can expand any vector ψ of the Hilbert space in the form

$$\psi = \sum_{n=1}^{\infty} c_n \xi_n, \tag{1}$$

i.e. the vector ψ has the components $c_1, c_2, \ldots c_n, \ldots$; these fully specify the vector ψ (since we are throughout referring to the basis

$\xi_1, \xi_2, \ldots \xi_n, \ldots$) and vice versa. In matrix algebra, we specify the vector ψ by its components $c_1, c_2, \ldots c_n, \ldots$ and write it as a *column vector*

$$\psi = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{2}$$

which is often abbreviated to $[c_n]$. We call (2) the matrix representation of ψ referred to the basis $\xi_1, \xi_2, \ldots \xi_n, \ldots$

Given two vectors $\psi = \sum_{n=1}^{\infty} c_n \xi_n$ and $\chi = \sum_{n=1}^{\infty} d_n \xi_n$, their inner product is $(\chi, \psi) = \sum_{n=1}^{\infty} \bar{d}_n c_n$. Anticipating matrix multiplication, with which we shall deal in section 22, we write this inner product in the form

$$(\chi, \psi) = [\bar{d}_1, \bar{d}_2, \ldots] \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \bar{d}_1 c_1 + \bar{d}_2 c_2 + \ldots = \sum_{=1}^{\infty} \bar{d}_n c_n, \tag{3}$$

i.e. we write the first factor in the inner product (χ, ψ) as a *row vector*, taking complex conjugates. We shall also refer to a row vector as a *transposed* vector and denote it by a prime, thus $[d_n]'$. Equation (3) can then be written

$$(\chi, \psi) = [\bar{d}_n]'[c_n] \tag{3a}$$

and the law of multiplication is to multiply corresponding components and to add these terms.

Let us consider the quantity corresponding to an operator A in matrix mechanics. Suppose

$$\chi = A\psi \tag{4}$$

and

$$\psi = \sum_{n=1}^{\infty} c_n \xi_n, \quad \chi = \sum_{n=1}^{\infty} d_n \xi_n,$$

i.e. $c_n = (\xi_n, \psi)$ and $d_n = (\xi_n, \chi)$. Substitution in (4) then gives

$$\sum_{m=1}^{\infty} d_m \xi_m = A \sum_{m=1}^{\infty} c_m \xi_m = \sum_{m=1}^{\infty} (A \xi_m) c_m$$

and taking the inner product with respect to ξ_n, since $(\xi_n, \xi_m) = \delta_{nm}$,

$$\begin{aligned} d_n &= \sum_{m=1}^{\infty} (\xi_n, A\xi_m) c_m \\ &= \sum_{m=1}^{\infty} A_{nm} c_m \quad (n = 1, 2, \dots) \end{aligned} \tag{5}$$

where we have introduced the abbreviation

$$A_{nm} = (\xi_n, A\xi_m). \tag{6}$$

(Note the order of the suffixes n and m, as in general $A_{nm} \neq A_{mn}$.) Equation (5) is the matrix equivalent of (4). It specifies how the nth component of χ is derived from the components of ψ. We see that a knowledge of the square array of numbers

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \dots\dots \\ A_{21} & A_{22} & A_{23} & \dots\dots\dots \\ A_{31} & A_{32} \dots & \dots & \dots\dots\dots \\ A_{41} & \dots & \dots & \dots\dots\dots \\ \dots & \dots & \dots & \dots\dots\dots \end{bmatrix} \tag{7}$$

is equivalent to a complete specification of the operator A; it is the matrix representation of A with respect to the basis $\xi_1, \xi_2, \dots \xi_n, \dots$ Given (7), we can, on account of (5), transform any vector ψ by means of A. We call the array (7) a *matrix operator*, or simply a *matrix*, and the A_{nm} occurring in (7), and defined by (6), *matrix elements*. We can write (5) in a different manner, using (7):

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \cdot \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \dots\dots \\ A_{21} & A_{22} & \dots\dots\dots \\ A_{31} & \dots & \dots\dots\dots \\ \dots & \dots & \dots\dots\dots \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \cdot \end{bmatrix}; \tag{5*}$$

this notation brings out particularly clearly that the matrix operator $[A_{nm}]$ transforms the vector $[c_n]$ into the vector $[d_n]$. The law of multiplication on the right-hand side of (5*) is as follows: the rth component of χ is obtained by multiplying the terms in the rth row of A_{nm} by the corresponding terms of ψ and adding these terms.

Above we have considered an infinitely dimensional function space whose vectors and matrices have infinitely many components and elements. But we can equally well apply the above formalism to a finite n-dimensional vector space whose vectors have n components and whose matrices have n^2 elements.

We can now express wave functions, inner products and operators in matrix form, enabling us to express many of our previous results in this language. Thus, the time-dependent Schrödinger equation (16.7),

$$i\hbar\frac{\partial\psi_t}{\partial t}=H\psi_t,$$

is replaced by the system of simultaneous differential equations

$$i\hbar\dot{c}_n(t)=\sum_{m=1}^{\infty}H_{nm}c_m(t)\quad(n=1,2,\ldots)\tag{8}$$

for the time-dependent expansion coefficients $c_n(t)$ of the wave function $\psi_t=\sum_{n=1}^{\infty}c_n(t)\xi_n$. Given the coefficients at $t=0$, $c_n(0)$, the system of equations (8) completely determines the $c_n(t)$, *i.e.* the way in which the individual components of the wave function vary with time.

To be able to use matrix mechanics generally we must consider the form which the most important properties of operators assume in matrix language. This we shall do in the next section.

22. Properties of matrix operators

We now derive the matrix form of some of the more important properties of the operators of quantum mechanics.

(a) *Matrix Multiplication*

To begin with, let us find the matrix corresponding to the product AB of two operators. For any vector φ, let $\psi=B\varphi$ and $\chi=A\psi$. $C=AB$ is the operator which transforms φ into $\chi:\chi=C\varphi=AB\varphi$. Suppose

$$\varphi=\sum_{n=1}^{\infty}b_n\xi_n,\quad\psi=\sum_{n=1}^{\infty}c_n\xi_n,\quad\chi=\sum_{n=1}^{\infty}d_n\xi_n.$$

We want the matrix representation of the operator $C=AB$, in terms of the matrices $[A_{nm}]$ and $[B_{nm}]$. Equation (21.5) gives at once

$$c_m=\sum_{l=1}^{\infty}B_{ml}b_l,\quad d_n=\sum_{m=1}^{\infty}A_{nm}c_m$$

and hence

$$d_n=\sum_{m=1}^{\infty}\sum_{l=1}^{\infty}A_{nm}B_{ml}b_l.\tag{1a}$$

But if $C_{nm}=(\xi_n,C\xi_m)$ is the matrix representation of the transformation C which transforms φ into $\chi=C\varphi$, we also have

$$d_n=\sum_{l=1}^{\infty}C_{nl}b_l.\tag{1b}$$

Now equations (1a) and (1b) are to hold for all vectors since C is simply defined as the product AB. Hence it follows that

$$C_{nl}=\sum_{m=1}^{\infty}A_{nm}B_{ml}, \quad (n, l=1, 2, \ldots) \tag{2}$$

or, using the complete matrix notation, we can write this

$$\begin{bmatrix} C_{11} & C_{12} & C_{13}\ldots \\ C_{21} & C_{22} & C_{23}\ldots \\ C_{31} & C_{32} & \ldots\ldots \\ \ldots & \ldots & \ldots\ldots \end{bmatrix}=\begin{bmatrix} A_{11} & A_{12} & A_{13}\ldots \\ A_{21} & A_{22} & A_{23}\ldots \\ A_{31} & A_{32} & \ldots\ldots \\ \ldots & \ldots & \ldots\ldots \end{bmatrix}\times\begin{bmatrix} B_{11} & B_{12} & B_{13}\ldots \\ B_{21} & B_{22} & B_{23}\ldots \\ B_{31} & B_{32} & \ldots\ldots \\ \ldots & \ldots & \ldots\ldots \end{bmatrix}. \tag{3}$$

Equation (2) states the law of matrix multiplication, which we have here deduced from the general transformation law of linear operators. We obtain the (1, 1) element of AB by multiplying the first row of A by the first column of B, term by term, adding these terms. Similarly we obtain the (m, n) element (*i.e.* the element in the mth row and nth column) of AB from the mth row of A and the nth column of B. But one must always multiply a row in the first factor by a column in the second factor and not in the reverse order.

As we would expect from the fact that matrices represent operators, matrix multiplication is in general not commutative,

$$[A_{nm}][B_{nm}]\neq[B_{nm}][A_{nm}].$$

We give two examples of this:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}=\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}=\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

For certain special matrices, multiplication is commutative. We define a diagonal matrix as being one in which all elements off the principal diagonal are zero:

$$[D_{nm}]=[d_n\delta_{nm}]=\begin{bmatrix} d_1 & 0 & 0\ldots\ldots \\ 0 & d_2 & 0\ \ldots\ldots \\ 0 & 0 & d_3\ \ldots\ldots \\ \ldots & \ldots & \ldots\ldots \end{bmatrix}. \tag{4}$$

In particular, the unit matrix, denoted by $\mathbf{I}$, is

$$[\mathbf{I}_{nm}]=[\delta_{nm}]=\begin{bmatrix} 1 & 0 & 0\ \ldots\ldots \\ 0 & 1 & 0\ \ldots\ldots \\ 0 & 0 & 1\ \ldots\ldots \\ \ldots & \ldots & \ldots\ldots \end{bmatrix}. \tag{5}$$

All matrices commute with the unit matrix, for

$$[A_{nm}][\delta_{nm}]=[\delta_{nm}][A_{nm}]=[A_{nm}],$$

and any two diagonal matrices commute and their product is also a diagonal matrix (*cf* Exercise V.1).

(*b*) *Adjoint Matrices*

If $A^\star$ is the adjoint of an operator A, then for any two vectors ψ and χ

$$(\chi, A\psi) = (A^\star\chi, \psi);$$

in particular for two basic vectors $\psi = \xi_m$, $\chi = \xi_n$,

$$A_{nm} = \overline{A}^\star{}_{mn} \tag{6}$$

i.e. the adjoint matrix is obtained by interchanging rows and columns (*i.e.* transposing the matrix) and taking complex conjugates. If A is self-adjoint,

$$A_{nm} = \overline{A}_{mn} \tag{7}$$

i.e. the diagonal elements $(m = n)$ are real and mirror elements (m, n) and (n, m) on either side of the principal diagonal are complex conjugates. A typical 2×2 Hermitian matrix is

$$\begin{bmatrix} x_1 & z \\ \bar{z} & x_2 \end{bmatrix} \tag{8}$$

where x_1 and x_2 are real and z may be complex.

(*c*) *Change of Basis*

Since the choice of basis is largely arbitrary, different representations being physically equally acceptable, we must now consider coordinate transformations from one complete orthonormal set of basic vectors $\xi_1, \xi_2, \ldots \xi_n, \ldots$ to another such set $\tilde{\xi}_1, \tilde{\xi}_2, \ldots \tilde{\xi}_n, \ldots$.

In Exercise IV.7 we saw how to generate such new bases from the original basis $\xi_1, \xi_2, \ldots \xi_n \ldots$: any unitary transformation U transforms $\xi_1, \xi_2, \ldots \xi_n, \ldots$ into a new complete orthonormal set of vectors $\tilde{\xi}_1, \tilde{\xi}_2, \ldots \tilde{\xi}_n, \ldots$ where

$$\tilde{\xi}_n = U\xi_n, \quad (n = 1, 2, \ldots)$$

which we can take as new basis.

Conversely, given any two complete orthonormal sets of vectors $\xi_1, \xi_2, \ldots \xi_n, \ldots$ and $\tilde{\xi}_1, \tilde{\xi}_2, \ldots \tilde{\xi}_n, \ldots$ we can define an operator U completely by the correspondence

$$\tilde{\xi}_n = U\xi_n, \quad (n = 1, 2, \ldots) \tag{9}$$

i.e. we define U as the operator which transforms ξ_n into $\tilde{\xi}_n$. Since the set of vectors $\xi_1, \xi_2, \ldots \xi_n, \ldots$ is complete, this also tells us the vector $\tilde{\psi}$ into which any vector ψ is transformed by U. For if $\psi = \sum_{n=1}^{\infty} c_n \xi_n$, then

$$\tilde{\psi} = U\psi = \sum_{n=1}^{\infty} c_n U \xi_n = \sum_{n=1}^{\infty} c_n \tilde{\xi}_n. \tag{10}$$

One easily sees that U possesses an inverse operator U^{-1} which is completely defined by the equations

$$\xi_n = U^{-1} \tilde{\xi}_n. \tag{11}$$

The operator U which we have defined is unitary; for, if $\psi = \sum_{n=1}^{\infty} c_n \xi_n$ and $\chi = \sum_{n=1}^{\infty} d_n \xi_n$, then

$$(\chi, \psi) = \sum_{n=1} \bar{d}_n c_n, \quad (U\chi, U\psi) = \sum_{m,\, n=1}^{\infty} \bar{d}_n c_m (\tilde{\xi}_n, \tilde{\xi}_m) = \sum_{n=1}^{\infty} \bar{d}_n c_n.$$

Hence equation (18.19) holds, $U^{-1} = U^{\star}$, and equation (11) becomes

$$\xi_n = U^{\star} \tilde{\xi}_n. \tag{12}$$

From the way in which we have introduced coordinate systems, it is clear that if we consider two vectors ψ and χ, whose column vectors, referred to the two bases, are $[c_n]$, $[d_n]$ and $[\tilde{c}_n]$, $[\tilde{d}_n]$ respectively, then

$$(\chi, \psi) = \sum_{n=1}^{\infty} \bar{d}_n c_n = \sum_{n=1}^{\infty} \bar{\tilde{d}}_n \tilde{c}_n;$$

the scalar product of two vectors ψ and χ is independent of the choice of basis to which we refer ψ and χ. This must of course be the case since scalar products give the physically observable quantities (expectation values, probabilities, *etc*).

To find the effect of a change of basis on the matrix representation of an operator A we consider the transformation of the vector ψ into $\chi = A\psi$ referred to two different bases. We are, of course, in both cases dealing with the same operator A and the same vectors ψ and χ, and only have different matrix representations of these. Referred to the basis $\xi_1, \xi_2, \ldots \xi_n, \ldots$, ψ and χ are given by $\psi = \sum_{n=1}^{\infty} c_n \xi_n$ and $\chi = \sum_{n=1}^{\infty} d_n \xi_n$, and the equation $\chi = A\psi$ becomes

$$d_n = \sum_{m=1}^{\infty} A_{nm} c_m \quad (n = 1, 2, \ldots) \tag{21.5}$$

where

$$A_{nm} = (\xi_n, A\xi_m). \tag{21.6}$$

Referred to the basis $\tilde{\xi}_1, \tilde{\xi}_2, \ldots \tilde{\xi}_n, \ldots$ the same vectors ψ and χ are

$\psi=\sum_{n=1}^{\infty}\tilde{c}_n\tilde{\xi}_n$ and $\chi=\sum_{n=1}^{\infty}\tilde{d}_n\tilde{\xi}_n$, and equations (21.5) become replaced by

$$\tilde{d}_n=\sum_{m=1}^{\infty}\tilde{A}_{nm}\tilde{c}_m, \quad (n=1, 2, \ldots), \tag{21.5a}$$

where $\tilde{A}_{nm}$ is the matrix representation of A referred to the new basis:

$$\tilde{A}_{nm}=(\tilde{\xi}_n, A\tilde{\xi}_m). \tag{21.6a}$$

From equations (21.6), (21.6a) and (9) we obtain

$$\tilde{A}_{nm}=(\tilde{\xi}_n, A\tilde{\xi}_m)=(U\xi_n, AU\xi_m)=(\xi_n, U^{\star}AU\xi_m)=(U^{\star}AU)_{nm},$$

or using the rule for matrix multiplication

$$\tilde{A}_{nm}=(U^{\star}AU)_{nm}=\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}U_{nr}{}^{\star}A_{rs}U_{sm}. \tag{13}$$

From (13) one can find the matrix representation of an operator with respect to a new basis.

The operator U we have so far considered can be thought of as a transformation (a kind of rotation) of the coordinate system in the function space, keeping the state vectors fixed. We obtain the same relative motion of coordinate system and state vectors if we consider the coordinate system as fixed and subject all state vectors to the inverse transformation $U^{\star}$. Since $U^{\star}$ is also a unitary operator (*cf* Exercise IV.5), we can generally think of a unitary operator U as a transformation which replaces all wave functions ψ by $\tilde{\psi}=U\psi$, leaving all inner products invariant, and replacing the matrix representation $[A_{nm}]$ of an operator A by $[\tilde{A}_{nm}]=[(UAU^{\star})_{nm}]$. For, if $\chi=A\psi$, then $\tilde{\chi}=\tilde{A}\tilde{\psi}$ implies $U\chi=\tilde{A}U\psi$, *i.e.*

$$\tilde{A}=UAU^{\star}. \tag{14}$$

(The order of the factors U and $U^{\star}$ here appears in the reverse order to that in (13), as we are rotating the vectors, instead of the coordinate system.)

(*d*) *Eigenvalue Problems*

We must now consider the form the eigenvalue problem takes in matrix mechanics. Let us, for the moment, assume we have already solved the eigenvalue problem of a Hermitian operator A, *i.e.* we know the eigenvalues $a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant \ldots$ and correspondingly a complete set of normed orthogonal eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots$ If we take this set of eigenfunctions as basis of a matrix representation of A, it follows from

$$(\varphi_n, A\varphi_m)=a_m(\varphi_n, \varphi_m)=a_m\delta_{nm} \tag{15}$$

that the matrix of A is a diagonal matrix with the eigenvalues of A as diagonal elements

$$[A_{nm}] = \begin{bmatrix} a_1 & 0 & 0 & \dots \\ 0 & a_2 & 0 & \dots \\ 0 & 0 & a_3 & \dots \\ 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}, \tag{16}$$

and the eigenfunction φ_1 has the matrix representation $[\delta_{1m}]$, *etc.* Though by a suitable choice of eigenfunctions we can always diagonalize $[A_{nm}]$, degenerate eigenfunctions belonging to the same eigenvalue will in general not be orthogonal. Hence the matrix A will not be diagonal, but consist of squares, containing non-zero elements, arranged along the principal diagonal. The eigenfunctions labelling the rows and columns of any square belong to the same eigenvalue. They span the corresponding manifold in Hilbert space. We illustrate this for the case $a_1 = a_2 \neq a_3 = a_4 \neq a_5 \neq \dots$,

$$[A_{nm}] = \left[\begin{array}{cc|cc|c} A_{11} & A_{12} & & & \\ A_{21} & A_{22} & & & 0 \\ \hline & & A_{33} & A_{34} & \\ & & A_{43} & A_{44} & \\ \hline & 0 & & & A_{55} \end{array}\right]. \tag{17}$$

To express the eigenvalue problem

$$A\varphi = a\varphi \tag{18}$$

in matrix form, we introduce a matrix representation in terms of some complete orthonormal basic set of vectors $\xi_1, \xi_2, \dots \xi_n \dots$ If we write the eigenfunction φ as $\varphi = \sum_{n=1}^{\infty} c_n \xi_n$, we must solve the infinite set of linear homogeneous equations

$$\sum_{m=1}^{\infty} A_{nm} c_m = a c_n \quad (n = 1, 2, \dots) \tag{19}$$

for the eigenvectors $[c_n]$. In general, this set of equations only has the trivial solution $[c_n] \equiv [0]$, and only for certain values of a, the eigenvalues, are there non-trivial solutions.

We shall not give a general treatment of the method of solution of a system of linear homogeneous equations, but shall only discuss the general features of the corresponding problem in an N-dimensional vector space $\mathfrak{R}_N$. This is the problem most frequently encountered in practice, the N-dimensional vector space being a given subspace of the infinitely dimensional function space. This more restricted eigenvalue problem leads to a finite number of equations for a finite number of unknowns. Its solution can always be put in the form of N orthonormal eigenvectors and corresponding eigenvalues.

We can formulate this eigenvalue problem as follows. Given the operator A, we can find its matrix representation $A_{nm}=(\xi_n, A\xi_m)$ with respect to an arbitrary basis $\xi_1, \xi_2, \ldots \xi_n, \ldots \xi_N$. We want to find a unitary transformation U which transforms this basis into the set of eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots \varphi_N$, which are initially unknown, to act as new basis:

$$\varphi_n = U\xi_n, \quad (n=1, 2, \ldots N). \tag{20}$$

Referred to this new basis, the matrix of A will have to be a diagonal matrix a, *i.e.*

$$U^{\star}AU=a,$$
$$AU=Ua,$$

or, in terms of matrix elements,

$$\begin{aligned}\sum_{m=1}^{N} A_{lm}U_{mn} &= \sum_{m=1}^{N} U_{lm}(a_m\delta_{mn}) \\ &= U_{ln}a_n \\ &= \sum_{m=1}^{N}\delta_{lm}U_{mn}a_n \quad (l, n=1, \ldots N),\end{aligned}$$

hence

$$\sum_{m=1}^{N}(A_{lm}-\delta_{lm}a_n)U_{mn}=0, \quad (l, n=1, \ldots N). \tag{21}$$

For a fixed value of n, equations (21) represent a system of N homogeneous linear equations. They contain the unknown parameter a_n and the N unknown quantities $U_{1n}, U_{2n}, \ldots U_{Nn}$. But since the equations (21) are homogeneous they cannot determine the absolute values of the U_{mn}, but only the $(N-1)$ ratios $U_{1n} : U_{2n} : \ldots . : U_{Nn}$. Hence equations (21) are a system of N equations for $(N-1)$ unknowns. Thus we have an over-determined problem and, in general, equations (21) will be inconsistent. To have a non-trivial solution, they must satisfy a consistency condition (which essentially states that one of the equations (21) is simply a linear combination of the remaining $(N-1)$ equations). This consistency condition is that the determinant of the coefficients of U_{mn} in (21) should vanish:

$$|A_{lm}-\delta_{lm}a_n|=0; \tag{22}$$

this is an equation of the Nth degree in a_n which has N generally distinct solutions which we shall denote by $a_1, a_2, \ldots a_N$. These are the eigenvalues of A. For each value a_n, one can solve the set of equations (21) to obtain the corresponding eigenfunction (20),

$$\varphi_n = U\xi_n = \sum_{m=1}^{\mathcal{N}} \xi_m(\xi_m, U\xi_n) = \sum_{m=1}^{\mathcal{N}} \xi_m U_{mn}, \quad (n=1, \ldots \mathcal{N}). \tag{23}$$

If the $\mathcal{N}$ roots of the 'secular equation', as equation (22) is called, are not all distinct, then we have a degenerate case. To solve the problem for a degenerate eigenvalue requires further analysis which we shall however omit. One finds that corresponding to an r-fold degenerate eigenvalue, the eigensolutions span an r-dimensional subspace of $\mathfrak{R}_{\mathcal{N}}$, and one can choose r orthonormal eigenfunctions to span this subspace.

With the aid of the above analysis all our previous results can be expressed in matrix form. One can easily show directly that a Hermitian operator has real eigenvalues and orthogonal eigenfunctions (*cf* Exercises V.2 and 3). Our results about commuting operators (section 19, theorem 2) take a particularly simple form. For if a set of observables possesses a complete set of simultaneous eigenfunctions, then, taking these eigenfunctions as basis, the observables of the set are all represented by diagonal matrices and these necessarily commute.

23. The equations of motion of operators

In section 16 we dealt with the time variation of a system. It is determined by the Schrödinger equation (16.7) or, in a particular representation, by the set of equations (21.8). In this description, known as the *Schrödinger picture*, the temporal variation of a system is given by a time-dependent wave function, the observables being described by time-independent operators.

In formulating matrix mechanics, Heisenberg adopted an alternative approach which we shall now develop for operators generally, rather than for a matrix representation.

This alternative mode of description, known as the *Heisenberg picture*, arises because the physically significant quantities are the inner products, of the form $(\chi, A\psi)$, as we have repeatedly stressed. We can then ascribe the time variation of such inner products to a time-dependence of the wave functions or of the operators.

In the Schrödinger picture, an observable $\mathfrak{A}$, which does not contain the time explicitly, is represented by a constant operator A_0, the time-dependence entering the wave functions ψ_t and χ_t through the Schrödinger equation (16.7) or (16.12) and (16.13)*.

In the Heisenberg picture, the states of the system are represented by time-independent wave functions ψ_0 and χ_0, and the time variation of the inner products arises from time-dependent operators A_t. The two modes of description agree at all times if

---

* We are throughout restricting ourselves to systems whose Hamiltonian functions do not contain the time explicitly.

$$\begin{aligned}(\chi_0, A_t\psi_0) &= (\chi_t, A_0\psi_t)\\ &= (T_t\chi_0, A_0 T_t\psi_0)\\ &= (\chi_0, T_t^{\star} A_0 T_t\psi_0),\end{aligned} \tag{1}$$

i.e. we define the operator A_t at time t in the Heisenberg picture by

$$A_t = T_t^{\star} A_0 T_t. \tag{2}$$

From (2) and (16.13) one easily obtains the differential equation which A_t satisfies:

$$\begin{aligned}i\hbar\frac{dA_t}{dt} &= i\hbar\frac{dT_t^{\star}}{dt}A_0 T_t + i\hbar T_t^{\star} A_0\frac{dT_t}{dt}\\ &= -HT_t^{\star}A_0T_t + T_t^{\star}A_0T_tH\\ &= -HA_t + A_tH\\ &= [A_t, H].\end{aligned} \tag{3}$$

(In deriving (3) we have used the fact that H commutes with T_t and $T_t^{\star}$ but not, in general, with A_0.) Equation (3) is the Heisenberg equation of motion of an operator which does not explicitly depend on the time ($\partial A_t/\partial t = 0$). One can easily generalize these results to the case when the operator A does depend explicitly on the time. We shall not discuss this case which is less important, but only state that the equation of motion (3) in this case becomes replaced by

$$i\hbar\frac{dA_t}{dt} = i\hbar\frac{\partial A_t}{\partial t} + [A_t, H]. \tag{4}$$

From these equations all our previous results (section 19) about constants of the motion can be derived. This one can verify directly. Alternatively, it follows from the fact that all our previous results were concerned with inner products and these, we have shown, agree in the two pictures. In section 19 we defined an operator A to be a constant of the motion if it does not contain the time explicitly and if it commutes with the Hamiltonian of the system. The reasonableness of this definition is seen from (4). For a constant of the motion all inner products $(\chi_0, A_t\psi_0)$ are strictly time-independent.

24. The Dirac notation for wave functions and operators

We now want to consider some developments in notation. In writing eigenfunctions we have so far written the corresponding eigenvalues as suffixes and have often explicitly written out the arguments of the wave functions. Since the emphasis is usually on the eigenvalues, it is often more convenient to write these as 'arguments' and omit the position coordinates. Thus we would write for the orbital angular momentum eigenfunctions (13.27) and (14.4), $\psi(E_{n,l}, l, m)$ and $\psi(l, m)$ instead of $\psi_{E_{n,l},l,m}$ and ψ_l^m, as we have done so far.

A similar change of notation is desirable for matrix elements. We shall, however, not write the matrix element $A_{mn} = (\xi_m, A\xi_n)$ as $A(m, n)$, but rather as

$$\langle m|A|n\rangle$$

and the matrix element $(\psi_{l'}^{m'}, H\psi_l^{m})$ say, as $\langle l', m'|H|l, m\rangle$.

This use of triangular brackets is due to Dirac who has developed a systematic notation along these lines. As this is frequently to be found in the literature we shall explain it briefly.

In Dirac's notation, a vector in Hilbert space, which so far we have denoted by $\varphi, \psi, \chi, \ldots$ is denoted by a *ket* vector $|\rangle$. Labels, in particular eigenvalues, distinguishing different ket vectors are written as arguments of the ket vectors. For example, we might write a set of vectors as $|1\rangle, |2\rangle, \ldots |n\rangle$, or the wave function (14.4) as $|l, m\rangle$.

If $\varphi \equiv |\rangle$, we denote the complex conjugate vector $\bar{\varphi}$ not by $\overline{|\rangle}$ in general, but by $\langle|$, which is called a *bra* vector. ($\langle\rangle$ is a *bra*-(c)-*ket*, hence the names of *bra* and *ket* vectors.) The ket and bra vectors correspond to the column and row vectors of matrix algebra. The scalar product of $\psi \equiv |b\rangle$ and $\varphi \equiv |a\rangle$ is written

$$(\varphi, \psi) = \langle a|b\rangle$$

and it follows that

$$\overline{\langle b|a\rangle} = \langle a|b\rangle.$$

If A is an operator, it is natural to write the scalar product of $|c\rangle = A|a\rangle$ and $|b\rangle$ as

$$(\psi, A\varphi) = \langle b|c\rangle = \langle b|A|a\rangle,$$

and generally for the matrix element between two states φ' and φ'' labelled by the sets of eigenvalues $\xi_1', \xi_2', \ldots \xi_n'$ and $\xi_1'', \xi_2'', \ldots \xi_n''$,

$$(\varphi'', A\varphi') = \langle \xi_1'', \xi_2'', \ldots \xi_n''|A|\xi_1', \xi_2', \ldots \xi_n'\rangle.$$

If A^* is the operator adjoint to A, then for any two vectors $|a\rangle$ and $|b\rangle$,

$$\langle a|A|b\rangle = \overline{\langle b|A^*|a\rangle},$$

being the statement analogous to equation (18.6).

EXERCISES V

V.1. Prove that two $N \times N$ diagonal matrices commute and their product is also a diagonal matrix.

V.2. Prove from equation (22.19) that the eigenvalues of a Hermitian matrix are real.

V.3. Prove from equation (22.19) that the eigenfunctions of a Hermitian matrix belonging to different eigenvalues are orthogonal.

V.4. The angular momentum operators for a particle with angular momentum $\hbar\sqrt{\left(\frac{1}{2}\left(\frac{1}{2}+1\right)\right)}$, spin $\frac{1}{2}$ as it is called, can be represented by the following matrices,

$$s_x=\frac{\hbar}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},\quad s_y=\frac{\hbar}{2}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix},\quad s_z=\frac{\hbar}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}.$$

Show that s_z and $\mathbf{s}^2=s_x{}^2+s_y{}^2+s_z{}^2$ are simultaneously observable and give their eigenvalues. Show that s_x, s_y, s_z do not commute with each other and obtain expressions for the commutators $[s_x, s_y]$, *etc.*

V.5. Solve the eigenvalue problem in the two-dimensional space $\mathfrak{R}_2$ of the operator A whose matrix, referred to the orthonormal vectors ξ_1 and ξ_2 as basis, is

$$[A_{mn}]=\begin{bmatrix}\lambda & \mu\\ \mu & \lambda\end{bmatrix}$$

[*i.e.* $\lambda=(\xi_1, A\xi_1)=(\xi_2, A\xi_2)$, $\mu=(\xi_1, A\xi_2)=(\xi_2, A\xi_1)$].

V.6. Derive the Heisenberg equation of motion (23.4) for an operator which depends explicitly on the time, the Hamiltonian of the system being time-independent.

V.7. Derive the matrix representation of the position coordinate x referred to the oscillator wave functions $\psi_n(x)$, equation (11.23), as basis.

VI

SYSTEMS OF MANY PARTICLES

SYSTEMS of many particles are of importance in many problems of atomic and nuclear physics. In these rather complicated problems a direct complete solution of the Schrödinger wave equation is not usually practicable. It is therefore important to gain as much information as possible by studying general aspects such as the angular momentum and symmetry properties of the system.

25. Angular momentum

In section 14 we considered the orbital angular momentum of a single particle, and we defined the operators l_x, l_y, l_z which, as is shown in Exercise IV.13, satisfy the commutation relations

$$[l_x, l_y] = i\hbar l_z \quad etc \tag{1}$$

$$[\mathbf{l}^2, l_x] = 0 \quad etc \tag{2}$$

where

$$\mathbf{l}^2 = l_x^2 + l_y^2 + l_z^2. \tag{3}$$

We found in section 14 that, for a single particle in a spherically symmetric field, the operators l_z, $\mathbf{l}^2$ and the Hamiltonian H form a complete set of commuting observables. The simultaneous eigenfunctions (13.27), in the notation of section 24, are

$$\psi(E_{n,l}, l, m) = R_{n,l}(r) Y_l^m(\theta, \varphi) \tag{4}$$

and have the eigenvalues $\hbar m$, $\hbar^2 l(l+1)$ and E_n, respectively, m and l taking on the values

$$m = -l, -l+1, \ldots, l-1, l; \quad l = 0, 1, 2, \ldots$$

Using the wave functions (4) as basis of a matrix representation we find that

$$\left.\begin{aligned}
\langle E_{n',l'}, l', m'|H|E_{n,l}, l, m\rangle &= E_{n,l}\, \delta_{nn'}\, \delta_{ll'}\, \delta_{mm'} \\
\langle E_{n',l'}, l', m'|\mathbf{l}^2|E_{n,l}, l, m\rangle &= \hbar^2 l(l+1)\, \delta_{nn'}\, \delta_{ll'}\, \delta_{mm'} \\
\langle E_{n',l'}, l', m'|l_z|E_{n,l}, l, m\rangle &= \hbar m \delta_{nn'}\, \delta_{ll'}\, \delta_{mm'}
\end{aligned}\right\} \tag{5}$$

and using equations (14.2) defining the angular momentum operators in polar coordinates one easily finds that l_x and l_y are not diagonal matrices in this representation.

We have derived equations (5) using the eigenfunctions (4) explicitly. However, equations analogous to (1), (2) and (3) hold for the angular momentum of any system, not just a single particle, and we shall see that a result analogous to (5) can be derived from them directly. The other point to note is that equations (5) give a representation by $(2l+1) \times (2l+1)$ matrices of the angular momentum for a given energy eigenvalue $E_{n,l}$. Since l can take on the value $0, 1, 2, \ldots$ this restricts us to 1×1, 3×3, $5 \times 5, \ldots$ matrices. But we can clearly write down similar representations with matrices having an even number of rows and columns. We shall see presently that this generalization to l values of half an odd integer is connected with the intrinsic angular momentum, the spin, of elementary particles.

To distinguish it from the orbital angular momentum **l** of a single particle we shall denote the general angular momentum vector by **M**. Assuming that **M** satisfies relations analogous to (1) and (2),

$$[M_x, M_y] = i\hbar M_z \quad etc \tag{6}$$

$$[M^2, M_z] = 0 \; etc \tag{7}$$

where M is the magnitude of the angular momentum vector,

$$M^2 = \mathbf{M}^2 = M_x{}^2 + M_y{}^2 + M_z{}^2, \tag{8}$$

we shall prove the following result which is sufficiently general to deal with particles with spin, systems of particles, *etc*:

Let Q, M^2, M_z form a complete set of commuting observables[*] for a system, which satisfy equations (6), (7), and assume that Q commutes not only with M_z but also with M_x and M_y,

$$[Q, M_x] = [Q, M_y] = [Q, M_z] = 0\dagger. \tag{9}$$

We shall see that the eigenvalues of M^2 are of the form

$$\hbar^2 n(n+1) \quad n = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots$$

Which of these eigenvalues actually occur in a particular system depends on the system. (Thus we saw that for the orbital angular momentum of a single particle we can only have $n = 0, 1, 2, \ldots$.) Furthermore, for a given eigenvalue q of Q, an eigenvalue $\hbar^2 n(n+1)$

---

[*] Q here stands as abbreviation for a set of commuting observables $Q_1, Q_2, \ldots$, if necessary. In general the Hamiltonian H will be one of these observables and it may happen that it is the only one, as in the orbital angular momentum case, above.

[†] This corresponds to the assumption in deriving (4) that the system is spherically symmetric. The Hamiltonian H is then invariant under a rotation of axes, *i.e.* it is independent of the angles θ and φ and hence commutes with l_x, l_y and l_z which, according to (14.2), only involve differentiations with respect to θ and φ. Conversely, if H commutes with l_x, l_y and l_z the Hamiltonian is spherically symmetric, *i.e.* independent of θ and φ. Thus, on account of (14.2), $l_z H\psi = H l_z \psi$ implies $\partial H/\partial\varphi = 0$, *i.e.* H is independent of φ.

of M^2 is $(2n+1)$-fold degenerate corresponding to $(2n+1)$ eigenstates $\Psi(q, n, m)$ of M_z with eigenvalues

$$\hbar m, \quad m = -n, -n+1, \ldots, n-1, n,$$

i.e.

$$Q\Psi(q, n, m) = q\Psi(q, n, m) \tag{10a}$$

$$M^2\Psi(q, n, m) = \hbar^2 n(n+1)\Psi(q, n, m), \quad (n = 0, \tfrac{1}{2}, 1, \ldots), \tag{10b}$$

$$M_z\Psi(q, n, m) = \hbar m\Psi(q, n, m), \quad (m = -n, \ldots n). \tag{10c}$$

We see that, as in the orbital angular momentum case, corresponding to a given pair of eigenvalues q, n there exists a $(2n+1)$-dimensional subspace in the Hilbert space of our system which is spanned by the $(2n+1)$ orthogonal eigenfunctions $\Psi(q, n, m)$ of M_z.

To prove these results from equations (6) to (9) we define the operators A_+ and A_- by

$$A_+ = M_x + iM_y, \quad A_- = M_x - iM_y, \tag{11}$$

so that $(A_+)^* = A_-$. One easily verifies the following commutation relations for these operators (*cf* Exercise VI.1)

$$[A_+, M_z] = -\hbar A_+ \tag{12a}$$

$$[A_-, M_z] = \hbar A_- \tag{12b}$$

$$[A_+, M^2] = [A_-, M^2] = 0 \tag{13}$$

and

$$A_+A_- = M^2 - M_z^2 + \hbar M_z \tag{14a}$$

$$A_-A_+ = M^2 - M_z^2 - \hbar M_z. \tag{14b}$$

Let $\Psi(q, \mu^2, \mu_z)$ be an eigenfunction of Q, M^2 and M_z with eigenvalues q, μ^2 and μ_z respectively, *i.e.*

$$Q\Psi(q, \mu^2, \mu_z) = q\Psi(q, \mu^2, \mu_z) \tag{15a}$$

$$M^2\Psi(q, \mu^2, \mu_z) = \mu^2\Psi(q, \mu^2, \mu_z) \tag{15b}$$

$$M_z\Psi(q, \mu^2, \mu_z) = \mu_z\Psi(q, \mu^2, \mu_z). \tag{15c}$$

Operating on (15c) with M_z and subtracting from (15b) we get

$$(M_x^2 + M_y^2)\Psi(q, \mu^2, \mu_z) = (\mu^2 - \mu_z^2)\Psi(q, \mu^2, \mu_z). \tag{16}$$

Since $(M_x^2 + M_y^2)$ is a positive definite operator it cannot have negative eigenvalues, *i.e.*

$$|\mu_z| < \mu \tag{17}$$

where $\mu > 0$ is the magnitude of $\mathbf{M}$ in the state $\Psi(q, \mu^2, \mu_z)$. It follows from (17) that the eigenvalues of M_z corresponding to a given eigenvalue of M^2 are bounded above and below.

Considering the wave functions*

$$A_{\pm}\Psi(q,\mu^2,\mu_z)$$

we easily prove, using equations (11) to (15), that

$$QA_{\pm}\Psi(q,\mu^2,\mu_z)=qA_{\pm}\Psi(q,\mu^2,\mu_z), \tag{18a}$$

$$M^2A_{\pm}\Psi(q,\mu^2,\mu_z)=\mu^2A_{\pm}\Psi(q,\mu^2,\mu_z), \tag{18b}$$

$$\begin{aligned} M_zA_{\pm}\Psi(q,\mu^2,\mu_z) &= A_{\pm}(M_z\pm\hbar)\Psi(q,\mu^2,\mu_z)\\ &=(\mu_z\pm\hbar)A_{\pm}\Psi(q,\mu^2,\mu_z), \end{aligned} \tag{18c}$$

i.e. $A_{\pm}\Psi(q,\mu^2,\mu_z)$ are eigenfunctions of Q, M^2 and M_z with the eigenvalues q, μ^2 and $(\mu_z\pm\hbar)$ respectively. Thus operating on $\Psi(q,\mu^2,\mu_z)$ with $A_{\pm}$ gives new simultaneous eigenfunctions, belonging to the *same* eigenvalues of Q and M^2 as $\Psi(q,\mu^2,\mu_z)$ itself, but to *different* eigenvalues of M_z.

Thus by repeatedly operating with A_+ on $\Psi(q,\mu^2,\mu_z)$ we generate a whole set of eigenfunctions of M_z belonging to the eigenvalues μ_z, $\mu_z+\hbar$, $\mu_z+2\hbar$, . . . and all belonging to the eigenvalues q and μ^2 of Q and M^2. Similarly, the operator A_- generates a set of eigenfunctions of M_z belonging to the eigenvalues μ_z, $\mu_z-\hbar$, $\mu_z-2\hbar$, . . . and to the eigenvalues q and μ^2 of Q and M^2. But on account of (17), $|\mu_z|$ cannot assume arbitrarily large values. Hence for a given total angular momentum μ^2, there must exist some maximum possible eigenvalue of M_z which we shall call $\hbar n$, so that A_+ operating on the corresponding eigenfunction $\Psi(q,\mu^2,\hbar n)$ does not give an eigenfunction of M_z corresponding to $\hbar(n+1)$, but vanishes identically:

$$A_+\Psi(q,\mu^2,\hbar n)\equiv 0. \tag{19}$$

If (19) holds identically then

$$A_-A_+\Psi(q,\mu^2,\hbar n)\equiv 0, \tag{20}$$

i.e. from (14b)

$$\mu^2-\hbar^2n^2-\hbar^2n=0$$

i.e

$$\mu^2=\hbar^2n(n+1). \tag{21}$$

Equation (21) defines the value of μ^2 for a given n and vice versa. Similarly there must exist a least eigenvalue $\hbar n'$ of M_z for which

$$A_-\Psi(q,\mu^2,\hbar n')\equiv 0 \tag{19a}$$

so that

$$A_+A_-\Psi(q,\mu^2,\hbar n')\equiv 0.$$

Using (14a) and (21) one finds

$$n(n+1)=n'(n'-1)$$

---

* In the following we shall use the symbols ± and ∓ as abbreviations for two alternatives. Thus $A_{\pm}\Psi(q,\mu^2,\mu_z)$ stands for the two wave functions $A_+\Psi(q,\mu^2,\mu_z)$ and $A_-\Psi(q,\mu^2,\mu_z)$ and an equation such as (18c) stands for the pair of equations obtained by taking the upper or lower signs throughout:

$$M_zA_+\Psi(q,\mu^2,\mu_z)=(\mu_z+\hbar)A_+\Psi(q,\mu^2,\mu_z)$$
$$M_zA_-\Psi(q,\mu^2,\mu_z)=(\mu_z-\hbar)A_-\Psi(q,\mu^2,\mu_z).$$

so that either $n'=n+1$ or $n'=-n$. The first alternative $n'=n+1$ is ruled out as $\hbar n$ was to be the largest eigenvalue of M_z, so that we must have $n'=-n$. We have thus proved that the manifold of eigenstates of Q and M^2 with the eigenvalues q and $\hbar^2 n(n+1)$ respectively is spanned by $(2n+1)$ eigenfunctions of M_z

$$\Psi(q, \hbar^2 n(n+1), \hbar m), \quad (m=-n, -n+1, \ldots n-1, n),$$

corresponding to the eigenvalues

$$\hbar m, \quad (m=-n, -n+1, \ldots n-1, n),$$

respectively. Since the dimensionality of this manifold is $(2n+1)$, $(2n+1)$ must be an integer, *i.e.* n can only have one of the values

$$n=0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \tfrac{5}{2}, \ldots$$

This completes the proof of equations (10), where we have for convenience written $\Psi(q, n, m)$ for the wave function $\Psi(q, \hbar^2 n(n+1), \hbar m)$ The corresponding matrix representation of Q, M^2 and M_z is

$$\langle q', n', m'|Q|q, n, m\rangle = q\delta_{qq'}\,\delta_{nn'}\,\delta_{mm'} \tag{22a}$$

$$\langle q', n', m'|M^2|q, n, m\rangle = \hbar^2 n(n+1)\delta_{qq'}\,\delta_{nn'}\,\delta_{mm'} \tag{22b}$$

$$\langle q', n', m'|M_z|q, n, m\rangle = \hbar m\delta_{qq'}\,\delta_{nn'}\,\delta_{mm'}. \tag{22c}$$

now complete,

Since the set of observables Q, M^2, M_z is complete, the eigenfunctions $\Psi(q, n, m)$ are non-degenerate, *i.e.* from (18c) $A_\pm\Psi(q, n, m)$ is a multiple of $\Psi(q, n, m\pm 1)$,

$$A_\pm\Psi(q, n, m) = \mathcal{N}_\pm(q, n, m)\Psi(q, n, m\pm 1). \tag{23}$$

$N_\pm$

The constants of proportionality $\mathcal{N}_\pm$ which may depend on q, n and m, are quite arbitrary in this equation. We can choose them in such a way that the $(2n+1)$ eigenfunctions $\Psi(q, n, n) \ldots, \Psi(q, n, -n)$, are all normed if we generate them by means of (23) starting with a normed eigenfunction $\Psi(q, n, m)$. One can easily show that this leads to the following values for $\mathcal{N}_\pm$, which depend on n and m only, so that we write $\mathcal{N}_\pm(n, m)$ instead of $\mathcal{N}_\pm(q, n, m)$:

$C_\pm = \hbar\sqrt{(j-m)(j+m+1)}$

$$\mathcal{N}_\pm(n, m) = \hbar\{(n\mp m)(n\pm m+1)\}^{\frac{1}{2}} \tag{24}$$

where we have taken an arbitrary phase factor of modulus unity equal to $+1$ (*cf* Exercise VI.2). From equations (23) and (24), we can then get the matrix elements of $A_\pm$ and hence of M_x and M_y:

$$\langle q, n, m+1|M_x|q, n, m\rangle = \tfrac{1}{2}\hbar\{(n-m)(n+m+1)\}^{\frac{1}{2}} \tag{25a}$$

$$\langle q, n, m+1|M_y|q, n, m\rangle = -\tfrac{1}{2}i\hbar\{(n-m)(n+m+1)\}^{\frac{1}{2}} \tag{25b}$$

$$\langle q, n, m-1|M_x|q, n, m\rangle = \tfrac{1}{2}\hbar\{(n+m)(n-m+1)\}^{\frac{1}{2}} \tag{25c}$$

$$\langle q, n, m-1|M_y|q, n, m\rangle = \tfrac{1}{2}i\hbar\{(n+m)(n-m+1)\}^{\frac{1}{2}}. \tag{25d}$$

All other matrix elements of M_x and M_y vanish.*

---

* We have omitted some of the rather tedious algebra leading to equations (24) and (25). The reader should easily be able to carry out the intermediate steps if he so desires. The more lazy reader may prefer to find it in Condon and Shortley[21], p. 48.

25. ANGULAR MOMENTUM

In section 14 we have shown that the eigenfunctions of the orbital angular momentum of a single particle are essentially the spherical harmonics $\Upsilon_l^m(\theta, \varphi)$, equation (13.20), θ and φ being the polar angles of the particle. The normalization and phase factor of the spherical harmonics were fixed by equations (13.19). In our present notation we can write the complete eigenfunctions as

$$\psi(q, n, m) = R_{q,n}\, \Upsilon_n^m(\theta, \varphi), \quad (n = 0, 1, 2 \ldots), \tag{26}$$

where $R_{q,n}$ is independent of m, and of θ and φ, and depends only on other coordinates and commuting observables not of interest to us. The equations (22) and (25) completely specify the eigenfunctions $\Psi(q, n, m)$, apart from a common phase factor. If for integral values of n, we choose this phase factor so that

$$\Psi(q, n, n) = \psi(q, n, n), \quad (n = 0, 1, \ldots), \tag{27}$$

then it can be shown that the eigenfunctions (26) are identical with the $\Psi(q, n, m)$ for all values of m:

$$\Psi(q, n, m) = \psi(q, n, m), \quad \begin{pmatrix} m = -n, \ldots, n \\ n = 0, 1, \ldots \end{pmatrix}. \tag{28}$$

$n = j$

We now see the reason for the choice of phase factors in our definition of spherical harmonics in section 13: the operators $A_\pm$ generate the spherical harmonics with these phase factors.

We have considered this question of the choice of phase factor in some detail as it is important to choose the phase factors in a consistent manner. This will become apparent when we consider the addition of angular momentum vectors in section 27.

As an example of the above theory we consider a system with angular momentum $\frac{1}{2}$. (It is usual to refer to a system as having angular momentum n or $\hbar n$ if the magnitude M of the angular momentum is $\hbar\sqrt{\{n(n+1)\}}$.) In section 26, where we shall deal with the interpretation of this case in detail, we shall see that it corresponds to an elementary particle of spin $\frac{1}{2}$, such as an electron or proton or neutron. The matrix representation (25) is in this case $(2n+1) = 2$-dimensional. If we write the eigenstates of M_z which have eigenvalues $\pm\hbar/2$ as

$$\Psi(\tfrac{1}{2}, \tfrac{1}{2}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \Psi(\tfrac{1}{2}, -\tfrac{1}{2}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{29}$$

the matrix representation of $\mathbf{M}$,

$$\mathbf{M} = \frac{\hbar}{2}\boldsymbol{\sigma}, \tag{30}$$

becomes

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{31}$$

where the *Pauli spin matrices* σ_x, σ_y, σ_z are the components of the vector $\boldsymbol{\sigma}$.

Using the Pauli matrices one can verify the general properties of angular momentum derived above, such as the commutation relations (6) which now are

$$\sigma_x\sigma_y - \sigma_y\sigma_x = 2i\sigma_z, \textit{ etc.} \tag{32}$$

One easily finds that

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I} \tag{33}$$

so that

$$M^2 = \hbar^2 \tfrac{1}{2}(\tfrac{1}{2}+1), \tag{34}$$

as required. The operators $A_\pm$ defined in (11) now become

$$A_\pm = \frac{\hbar}{2}\left\{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \pm i \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\right\}$$

i.e.

$$A_+ = \frac{\hbar}{2}\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad A_- = \frac{\hbar}{2}\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}, \tag{35}$$

(as required $(A_\pm)^* = A_\mp$). We can check the process used above for generating eigenfunctions by means of the operators $A_\pm$. Thus

$$A_+\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$A_-\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

agreeing with equations (19) and (19a). Similarly

$$A_+\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \hbar\begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$A_-\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{\hbar}{2}\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \hbar\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

in agreement with (23) and (24). Another useful property of the Pauli matrices is that they *anticommute*, *i.e.*

$$\sigma_x\sigma_y + \sigma_y\sigma_x = 0, \textit{ etc.} \tag{36}$$

26. Spin

In discussing spin we shall assume that the reader is familiar with the experimental facts in spectroscopy, such as the anomalous Zeeman effect, which led Goudsmit and Uhlenbeck to the hypothesis of the electron spin. We shall only state briefly the main facts which led to this hypothesis and shall then consider its interpretation and mathematical formulation.

26. SPIN

Being electrically charged, an electron interacts with a magnetic field. For a homogeneous field H_z in the z-direction, the interaction energy, which enters the Schrödinger equation as an additional term, is

$$\frac{e}{2\mu c} l_z H_z \tag{1}$$

where l_z is the z-component of the orbital angular momentum of the electron ($-e$ = electronic charge, μ = electronic mass, c = velocity of light). We can interpret (1) by ascribing a magnetic moment

$$\frac{-e}{2\mu c}\mathbf{l} \tag{2}$$

to the electron. One obtains this magnetic moment classically by thinking of the electron in its orbit as defining a magnetic shell.

The effect of the interaction term (1) is to cause a shifting in the atomic energy levels which results in a change in the line spectrum when an external magnetic field is applied to an atom. This is the normal Zeeman effect which we shall consider in Chapter VII.

The observed spectrum of an atom in a magnetic field, known as the anomalous Zeeman effect, is much more complicated and cannot be explained in the above way. Many other spectroscopic phenomena such as the multiplicity of many spectral lines, which on the above theory one would have expected to be single lines, also defy any simple explanation without the introduction of a new hypothesis.

The hypothesis which resolved many of these difficulties, and which is essential to an understanding of many atomic and nuclear properties, is that of the electron *spin*. We assume that associated with the electron is an intrinsic angular momentum or spin $\frac{1}{2}$, *i.e.* the magnitude s of the spin vector $\mathbf{s}$ is given by

$$s^2 = \hbar^2 \tfrac{1}{2}(\tfrac{1}{2}+1).$$

$j = 1/2$

The spin thus represents an example of an angular momentum $\frac{1}{2}$ treated at the end of section 25, and we can write the spin vector $\mathbf{s}$ in terms of the Pauli vector $\boldsymbol{\sigma}$, (25.30),

$$\mathbf{s} = \tfrac{1}{2}\hbar\boldsymbol{\sigma}. \tag{3}$$

Associated with the spin $\frac{1}{2}$ is an intrinsic magnetic moment

$$\frac{-e}{\mu c}\mathbf{s}. \tag{4}$$

The theory of the spin of the electron based on the above postulates of spin and magnetic moment is due to Pauli. That quantum mechanics was successful in explaining many atomic properties without the introduction of spin is due to the fact that, in atomic systems, the spin interaction energy is a small correction term in the Hamiltonian of the system: to a first approximation the effect of spin can be neglected.

Pauli's theory is a non-relativistic theory. It is one of its peculiarities that one must postulate the intrinsic magnetic moment (4) to obtain agreement with experiment. For, by comparison with (2), we should have expected an intrinsic magnetic moment of half the magnitude, $-e\mathbf{s}/2\mu c$. In Chapter X we shall consider the relativistic theory of the electron due to Dirac. This theory predicts the correct spin and magnetic moment of the electron. In this chapter we shall only consider the non-relativistic theory due to Pauli, which is of course a limiting form of Dirac's theory. For many applications one only requires the non-relativistic theory since relativistic effects are negligible.

Apart from electrons, there are other elementary particles which possess spin $\frac{1}{2}$, for example protons and neutrons. Though we shall for definiteness often speak of electrons, many of our considerations will apply to these other cases as well. The limiting criterion will in all cases be that a non-relativistic theory should be adequate, *i.e.* that the energies of importance in the phenomena considered should be small compared to the rest energy of the elementary particles involved. For electrons this rest energy is 0·5 MeV, for protons and neutrons it is 940 MeV. As the binding energies of atomic electrons are of the order of a few electron volts and the binding energies of protons and neutrons in nuclei are of the order of a few MeV, the non-relativistic theories are applicable in a very wide field of atomic, molecular and nuclear physics.

On account of the spin, the electron now has four degrees of freedom. To specify its state we must supplement the position coordinates x, y, z by one other observable. We can take for this the z-component s_z of the spin angular momentum. The electron wave function is now $\Phi(\mathbf{r}, s_z)$. Whereas $\mathbf{r}$ can assume all values, s_z only has the two eigenvalues $\pm\hbar/2$,

$$s_z\Phi(\mathbf{r}, \pm\hbar/2) = \pm\tfrac{1}{2}\hbar\Phi(\mathbf{r}, \pm\hbar/2); \tag{5}$$

$\Phi(\mathbf{r}, \pm\hbar/2)$ represent states of the electron in which the z-component of its spin has the exact values $\pm\hbar/2$,* we say, the spin points in the direction of the $\pm z$-axis. In general we have a linear superposition of these states, and there exist definite probabilities of obtaining the values $\pm\hbar/2$ as a result of measuring s_z. We can think of $\Phi(\mathbf{r}, s_z)$ as a wave function with two components Φ_1 and Φ_2, and write

$$\left.\begin{aligned}\Phi_1 &= \Phi(\mathbf{r}, +\hbar/2) = \varphi_1(\mathbf{r})\alpha = \varphi_1(\mathbf{r})\begin{bmatrix}1\\0\end{bmatrix}\\ \Phi_2 &= \Phi(\mathbf{r}, -\hbar/2) = \varphi_2(\mathbf{r})\beta = \varphi_2(\mathbf{r})\begin{bmatrix}0\\1\end{bmatrix}.\end{aligned}\right\} \begin{array}{l}\text{spin up}\\ \text{spin down}\end{array} \tag{6}$$

---

* One often measures spin and angular momentum in units of $\hbar$ and refers to these eigenvalues as $\pm\frac{1}{2}$. Some authors use σ_z instead of s_z as fourth observable with corresponding changes in notation.

In (6) we explicitly introduced the spin eigenfunctions (25.29)

$$\alpha = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \beta = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{7}$$

of section 25, and the wave functions $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{r})$ in (6) are now ordinary spatial wave functions which are independent of spin. We can combine equations (6) and put

$$\Phi = \Phi_1 + \Phi_2 = \varphi_1 \alpha + \varphi_2 \beta = \varphi_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \varphi_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix}. \tag{8}$$

We have here introduced several alternative notations which all express the same facts. It is important to be familiar with all of them as they are all frequently used. From (6) we see that

$$s_z \Phi(\mathbf{r}, +\hbar/2) = s_z \varphi_1(\mathbf{r}) \alpha = +\tfrac{1}{2}\hbar \varphi_1(\mathbf{r}) \alpha$$
$$s_z \Phi(\mathbf{r}, -\hbar/2) = s_z \varphi_2(\mathbf{r}) \beta = -\tfrac{1}{2}\hbar \varphi_2(\mathbf{r}) \beta,$$

i.e. $\varphi_1(\mathbf{r})\alpha$ and $\varphi_2(\mathbf{r})\beta$ are the parts of the wave function representing spin states with $s_z = \pm\hbar/2$. The probability of finding an electron in the volume V with z-component of spin $\pm\hbar/2$ is

$$\int_V |\varphi_i(\mathbf{r})|^2 d\tau$$

with $i = 1, 2$ respectively. The probability of finding an electron in a volume V regardless of its spin state is

$$\int_V \{|\varphi_1(\mathbf{r})|^2 + |\varphi_2(\mathbf{r})|^2\}\, d\tau,$$

and for the normalization of the wave function we obtain

$$\int_{\substack{\text{all} \\ \text{space}}} \{|\varphi_1(\mathbf{r})|^2 + |\varphi_2(\mathbf{r})|^2\}\, d\tau = 1. \tag{9}$$

The combination of spin and spatial wave functions allows a particularly lucid formulation in terms of the ideas of vector spaces considered in Chapter I. We described particles without spin by ordinary wave functions $\varphi(\mathbf{r})$, $\psi(\mathbf{r})$, $\chi(\mathbf{r})$, . . . in a function space $\mathfrak{H}$ spanned by a complete orthonormal set of functions $\xi_1(\mathbf{r})$, $\xi_2(\mathbf{r})$, . . . $\xi_n(\mathbf{r})$, . . . We can describe the spin state of a particle by a vector in a two-dimensional spin space $\mathfrak{S}$. If we span $\mathfrak{S}$ by the basic vectors α and β such a spin vector has the form $a\alpha + b\beta$. A general electron wave function, equation (8), can then be thought of as a vector in the product space $\mathfrak{H} \times \mathfrak{S}$ formed from $\mathfrak{H}$ and $\mathfrak{S}$. The product space $\mathfrak{H} \times \mathfrak{S}$ consists of the products of all wave functions of $\mathfrak{H}$ by all wave functions of $\mathfrak{S}$; it is spanned by the complete orthonormal set $\xi_1(\mathbf{r})\alpha$, $\xi_1(\mathbf{r})\beta$, . . . $\xi_n(\mathbf{r})\alpha$, $\xi_n(\mathbf{r})\beta$, . . . If $\varphi_1(\mathbf{r}) = \sum_n c_n \xi_n$ and $\varphi_2(\mathbf{r}) = \sum_n d_n \xi_n$, we can write

$$\Phi = \varphi_1(\mathbf{r})\alpha + \varphi_2(\mathbf{r})\beta = \sum_n \xi_n(\mathbf{r}) \{c_n \alpha + d_n \beta\}$$

and we can interpret Φ either as a vector in spin space $\mathfrak{S}$, the amplitudes of the two spin states (*i.e.* the amounts of spin $+\frac{1}{2}$ and $-\frac{1}{2}$) being functions of position, or as a vector in $\mathfrak{H}$, with each orbital wave function $\xi_n(\mathbf{r})$ which occurs in Φ being a superposition of the two spin states.

We shall usually find it most convenient to introduce the representation with α and β as basic vectors in the spin space $\mathfrak{S}$. In $\mathfrak{H}$ we shall usually prefer the coordinate-free formulation so as to be able to treat each problem in the manner best suited to it. This leads to a difference in the appearance of quantities referring to spin and spatial coordinates. The former occur naturally in matrix form, the latter in coordinate-free form.

We define the inner product of two vectors

$$\Psi=\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}\qquad \Phi=\begin{bmatrix}\varphi_1\\ \varphi_2\end{bmatrix}$$

in the product space $\mathfrak{H}\times\mathfrak{S}$ by

$$(\Phi,\Psi)=\int\left\{[\bar{\varphi}_1,\bar{\varphi}_2]\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix}\right\}d\tau=\int\{\bar{\varphi}_1\psi_1+\bar{\varphi}_2\psi_2\}\,d\tau$$
$$=(\varphi_1,\psi_1)+(\varphi_2,\psi_2) \tag{10}$$

i.e. the scalar product in $\mathfrak{H}\times\mathfrak{S}$ is formed by adding the scalar products of the corresponding components of Φ and Ψ in $\mathfrak{H}$, in agreement with (9).

Operators in the product space $\mathfrak{H}\times\mathfrak{S}$ can now be subdivided into operators which operate in $\mathfrak{S}$ only, in $\mathfrak{H}$ only and combinations of these two which operate in both spaces.

The operators in spin space are 2×2 matrices with constant coefficients

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}\varphi_1\\ \varphi_2\end{bmatrix}=\begin{bmatrix}a\varphi_1+b\varphi_2\\ c\varphi_1+d\varphi_2\end{bmatrix}$$

i.e. such operators transform each component φ_i into a linear combination of φ_1 and φ_2. If the operator represents an observable it must be Hermitian, *i.e.* a and d are real and $b=\bar{c}$. The Pauli matrices are Hermitian as required.

Apart from these operators in $\mathfrak{S}$ we have operators in $\mathfrak{H}$ which are functions of the coordinates and momenta just as in the spin-free theory. These operators are now understood to be applied to both components of the wave function Φ, *e.g.*

$$F(\mathbf{r},\,-i\hbar\,\mathrm{grad})\begin{bmatrix}\varphi_1\\ \varphi_2\end{bmatrix}=\begin{bmatrix}F(\mathbf{r},\,-i\hbar\,\mathrm{grad})\varphi_1\\ F(\mathbf{r},\,-i\hbar\,\mathrm{grad})\varphi_2\end{bmatrix};$$

more accurately we should write such operators as multiples of the unit matrix in spin space, *i.e.* if we simply write F for $F(\mathbf{r},\,-i\hbar\,\mathrm{grad})$,

$$F=F\mathbf{I}=F\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}F & 0\\ 0 & F\end{bmatrix}. \tag{11}$$

From (11) it follows that an operator in $\mathfrak{H}$ only, commutes with an operator in $\mathfrak{S}$ only, since any operator of $\mathfrak{H}$ is a multiple of the unit matrix in the spin space, and any operator in $\mathfrak{S}$ contains only constant matrix elements and hence always commutes with any operator $F(\mathbf{r}, -i\hbar\,\mathrm{grad})$. Thus the operators referring to the orbital angular momentum of an electron, $l_x, l_y, l_z, \mathbf{l}^2$, always commute with operators referring to the electron spin $s_x, s_y, s_z, \mathbf{s}^2$,

$$[s_i, l_j]=0 \quad (i, j=x, y, z) \quad etc. \tag{12}$$

We can also define operators which operate in both $\mathfrak{H}$ and $\mathfrak{S}$, *e.g.*

$$\begin{bmatrix} F(\mathbf{r}, -i\hbar\,\mathrm{grad}), & G(\mathbf{r}, -i\hbar\,\mathrm{grad}) \\ H(\mathbf{r}, -i\hbar\,\mathrm{grad}), & I(\mathbf{r}, -i\hbar\,\mathrm{grad}) \end{bmatrix},$$

but we shall not require these.

With the above definitions of operators, inner products, *etc*, we obtain expectation values, probabilities, *etc* in exactly the same way as for particles without spin. Consider the wave function

$$\Phi=c_1\varphi_1(\mathbf{r})\,\alpha+c_2\varphi_2(\mathbf{r})\,\beta$$

with $(\varphi_1, \varphi_1)=(\varphi_2, \varphi_2)=1$. The normalization of Φ then requires

$$|c_1|^2+|c_2|^2=1.$$

The probability of the electron described by Φ having a z-component of spin $\pm\frac{1}{2}$ is $|c_1|^2$, $|c_2|^2$ respectively. The expectation value of its x-component of momentum is

$$\begin{aligned}\langle p_x\rangle=(\Phi, p_x\Phi)&=\left((c_1\varphi_1\alpha+c_2\varphi_2\beta),\; -i\hbar\frac{\partial}{\partial x}(c_1\varphi_1\alpha+c_2\varphi_2\beta)\right)\\ &=|c_1|^2(\varphi_1, p_x\varphi_1)+|c_2|^2(\varphi_2, p_x\varphi_2),\end{aligned}$$

i.e. it is the mean of the expectation values of p_x in the states φ_1 and φ_2, suitably weighted.

As second example we consider an atomic electron whose Hamiltonian function can be written

$$H=H_0+H_s \tag{13}$$

where H_0 is a term independent of spin which is spherically symmetric and H_s represents a spin-dependent term which, in general, is a small correction compared with H_0. Since H_0 is spherically symmetric, it follows from section 25 that H_0 commutes with the orbital angular momentum operators

$$[H_0, l_x]=0 \quad etc. \tag{14}$$

In the spin-free theory we neglect spin completely. H_0 then represents the complete Hamiltonian of the system and, on account of (14), the total orbital angular momentum and its z-component are constants of the motion. The energy levels of the system, *i.e.* the eigenvalues $E_{n,l}$ of H_0 are $(2l+1)$-fold degenerate, as follows from equations (25.10).

As the second approximation we allow for electron spin, but still ignore the spin term H_s in the Hamiltonian. Each eigenfunction $\psi(E_{n,l}, l, m)$ of H_0 gives us two eigenfunctions $\psi(E_{n,l}, l, m)\alpha$ and $\psi(E_{n,l}, l, m)\beta$ belonging to the *same* energy level $E_{n,l}$, whose degeneracy thus becomes doubled to $2(2l+1)$. In addition to (14) we now have

$$[H_0, s_x] = [H_0, s_y] = [H_0, s_z] = [H_0, \mathbf{s}^2] = 0 \qquad (15)$$

since H_0 is a multiple of the unit matrix in spin space, *i.e.* we can take the operators H_0, $\mathbf{l}^2$, l_z and s_z, which are constants of the motion, as a complete set specifying the system*.

Finally, we consider the accurate Hamiltonian function (13). The effect of the spin term H_s is to shift and in general split the energy levels $E_{n,l}$, as different spin states have different interaction energies. In general, the orbital and spin angular momenta are no longer constants of motion. Instead, the total angular momentum of the electron $\mathbf{j} = \mathbf{l} + \mathbf{s}$ commutes with the Hamilton H. The state of the system is now specified by the eigenvalues of $\mathbf{j}^2$ and j_z, each such state being a superposition of eigenstates of the orbital and spin angular momentum operators. To investigate this problem more fully we must be able to add angular momenta in quantum mechanics. We consider this problem in section 27.

27. Addition of angular momenta

We must next consider how the angular momentum of a composite system is related to the angular momenta of its constituents, *i.e.* the addition of angular momenta. This problem is of great practical importance in nuclear and atomic physics where we often deal with systems of many particles (protons, neutrons, electrons), and it is the resultant angular momentum of the nucleus or atom which is of direct physical interest.

In section 26 we considered in detail the combination of spatial and spin wave functions of a particle and found that the combined system is described in the product space $\mathfrak{H} \times \mathfrak{S}$, which is a natural generalization of the spaces $\mathfrak{H}$ and $\mathfrak{S}$ appropriate for the description of spatial or spin properties by themselves.

Before considering addition of angular momenta, we must deal with this problem of combining two systems Σ_1 and Σ_2 generally. The situation is analogous to that of section 26. By itself, each system Σ_k $(k=1, 2)$ is described in a function space $\mathfrak{F}_k$, spanned by a complete set of orthonormal vectors $\xi_1^{(k)}, \xi_2^{(k)}, \ldots \xi_n^{(k)}, \ldots$ The combined system Σ is described in the product space $\mathfrak{F} = \mathfrak{F}_1 \times \mathfrak{F}_2$, spanned by the complete orthonormal set $\xi_m^{(1)}\, \xi_n^{(2)}$ $(m, n = 1, 2, \ldots)$. The scalar products in $\mathfrak{F}$ are defined if we define them for the basic vectors $\xi_m^{(1)}\,\xi_n^{(2)}$ in terms of which any vector of $\mathfrak{F}$ can be expanded. We define this scalar product as

---

* For a single electron $\mathbf{s}^2$ is of course always a constant of the motion, with the value $\frac{3}{4}\hbar^2$, since both α and β and hence any vector of $\mathfrak{H} \times \mathfrak{S}$ are eigenfunctions of $\mathbf{s}^2$ with this eigenvalue.

$$(\xi_m^{(1)}\,\xi_n^{(2)},\ \xi_\mu^{(1)}\,\xi_\nu^{(2)}) = (\xi_m^{(1)},\ \xi_\mu^{(1)})(\xi_n^{(2)},\ \xi_\nu^{(2)}) = \delta_{m\mu}\delta_{n\nu}. \qquad (1)$$

Operators which refer to one system only, say to Σ_1, operate only in $\mathfrak{F}_1$; as far as $\mathfrak{F}_2$ is concerned they are unit operators. But operators representing interactions between the two systems are of a more general type and operate in both $\mathfrak{F}_1$ and $\mathfrak{F}_2$. The interaction between the spin $\mathbf{s}$ and the orbital angular momentum $\mathbf{l}$ of an atomic electron, which can be shown to be proportional to $\mathbf{ls}$, represents such an operator in $\mathfrak{H} \times \mathfrak{S}$.

We obtain further examples of product spaces by considering the spatial or spin wave functions of two electrons. In the first case we operate in the product space $\mathfrak{H} = \mathfrak{H}_1 \times \mathfrak{H}_2$, $\mathfrak{H}_k$ being the Hilbert space of the spatial wave functions $\varphi(\mathbf{r}_k)$ of the kth electron. The definition (1) of the inner product in $\mathfrak{H}_1 \times \mathfrak{H}_2$ is consistent with our previous definition, since

$$\iint \overline{\varphi(\mathbf{r}_1)\psi(\mathbf{r}_2)} f(\mathbf{r}_1) g(\mathbf{r}_2)\, d\tau_1 d\tau_2 = \int \overline{\varphi(\mathbf{r}_1)} f(\mathbf{r}_1)\, d\tau_1 \int \overline{\psi(\mathbf{r}_2)} g(\mathbf{r}_2)\, d\tau_2.$$

Operators referring to particle 1 only (*e.g.* $\mathbf{r}_1$ or $-i\hbar\,\mathrm{grad}_1$) are unit operators as far as particle 2 is concerned, *e.g.*

$$(f(\mathbf{r}_1)g(\mathbf{r}_2),\ \mathbf{r}_1\varphi(\mathbf{r}_1)\psi(\mathbf{r}_2)) \ = \ (f(\mathbf{r}_1),\ \mathbf{r}_1\varphi(\mathbf{r}_1))(g(\mathbf{r}_2),\ \psi(\mathbf{r}_2));$$

a typical interaction operator would be $e^2/(|\mathbf{r}_1 - \mathbf{r}_2|)$ representing the Coulomb repulsion between the electrons. The spin space $\mathfrak{S}_1 \times \mathfrak{S}_2$ of two electrons, spanned by $\alpha_1\alpha_2$, $\alpha_1\beta_2$, $\beta_1\alpha_2$, $\beta_1\beta_2$, represents a second example of a product space.

After these preliminaries, we return to consider the addition of the angular momentum vectors of two systems Σ_k, $(k = 1, 2)$, described in function spaces $\mathfrak{F}_k$. For each system by itself we define an angular momentum vector $\mathbf{M}(k)$ and the related operators

$$A_\pm(k) = M_x(k) \pm iM_y(k) \quad (k = 1, 2) \qquad (2)$$

as in section 25, and we assume that for each system commutation relations analogous to equations (25.6–8) and (25.11–14) hold. Since the operators $\mathbf{M}(1)$ and $\mathbf{M}(2)$ refer to two different systems, they commute, *i.e.*

$$[M_r(1),\ M_s(2)] = 0 \qquad (3)$$

where $M_r(1)$ stands for any component of $\mathbf{M}(1)$ and similarly for $M_s(2)$.

We define the total or resultant angular momentum vector of the combined system Σ by

$$\mathbf{M} = \mathbf{M}(1) + \mathbf{M}(2) \qquad (4)$$

with analogous definitions for the components and for $A_\pm$. On account of (3) one can easily prove that the total angular momentum operator

M satisfies the same commutation relations (25.6–8) and (25.11–14) as **M**(1) and **M**(2) separately (*cf* Exercise VI.7).

As in section 25, the angular momentum operators $M^2(k)$, $M_z(k)$ have in general to be supplemented by a set of operators Q_k to form a complete set of commuting observables for the system Σ_k. For the combined system, we could take the set of commuting operators Q_1, Q_2, $M^2(1)$, $M^2(2)$, $M_z(1)$ and $M_z(2)$ as a complete set. But since we are interested in the resultant angular momentum of Σ we prefer to replace the observables $M_z(1)$ and $M_z(2)$ by M^2 and M_z, and to take the complete set of observables Q_1, Q_2, $M^2(1)$, $M^2(2)$, M^2 and M_z. Since we are only interested in the angular momentum properties of the system we shall for simplicity omit the operators Q_1 and Q_2 and their eigenvalues, assuming that we are always dealing with the same set of eigenvalues q_1 and q_2. The state of the combined system can then be fully specified by either $M^2(1)$, $M^2(2)$, $M_z(1)$, $M_z(2)$, or $M^2(1)$, $M^2(2)$, M^2 and M_z.

Since the total angular momentum **M** satisfies all the conditions stipulated in section 25, the main result of that section holds: M^2 can only have the eigenvalues $\hbar^2 n(n+1)$, $(n=0, \frac{1}{2}, 1, \ldots)$ and these are $(2n+1)$-fold degenerate corresponding to M_z having the eigenvalues $\hbar m$, $(m=-n, -n+1, \ldots n-1, n)$.

Suppose that $M(k)$ has the value n_k for the system Σ_k, $(k=1, 2)$, what are the allowed values of the resultant angular momentum M and how are the corresponding eigenfunctions derived from those of the constituent systems?* We shall not give the complete treatment of this problem which is straightforward but rather tedious. It can be found in many standard text-books (*e.g.* Condon and Shortley[21] or van der Waerden[23]). We shall only give enough of the theory to make the results appear plausible and we shall illustrate their application to particular problems.

Suppose the angular momentum $M(k)$ of Σ_k has the value n_k which is $(2n_k+1)$-fold degenerate. The wave function of Σ_k is then a vector lying in the $(2n_k+1)$-dimensional manifold $\mathfrak{M}_{n_k}(k)$ of $\mathfrak{F}_k$ spanned by $\Psi_k(n_k, -n_k), \ldots \Psi_k(n_k, n_k)$, using the notation of section 25 but omitting the eigenvalues q_k. The possible states of the combined system then lie in the $(2n_1+1)(2n_2+1)$-dimensional manifold $\mathfrak{M}_{n_1}(1) \times \mathfrak{M}_{n_2}(2)$ spanned by the product wave functions

$$\Psi_1(n_1, m_1)\Psi_2(n_2, m_2) \quad (m_k = -n_k, \ldots n_k). \tag{5}$$

Corresponding to this manifold, the total angular momentum M^2 assumes each of the values $\hbar^2 n(n+1)$, with

$$n = n_1+n_2,\ n_1+n_2-1, \ldots n_1-n_2+1,\ n_1-n_2, \tag{6}$$

---

* We shall again for brevity say that $M(k)$ has the value n_k when we mean that the system is in an eigenstate of $M^2(k)$ corresponding to the eigenvalue $\hbar^2 n_k(n_k+1)$.

where for definiteness we have assumed that $n_1 \geqslant n_2$. Each eigenvalue $\hbar^2 n(n+1)$ is $(2n+1)$-fold degenerate; the corresponding manifold $\mathfrak{M}_n$ is spanned by the $(2n+1)$ eigenfunctions of M_z. We can express this decomposition of the product manifold $\mathfrak{M}_{n_1}(1) \times \mathfrak{M}_{n_2}(2)$ into manifolds $\mathfrak{M}_n$, which are mutually orthogonal, each of which corresponds to one definite eigenvalue of M, symbolically by the equation

$$\mathfrak{M}_{n_1}(1) \times \mathfrak{M}_{n_2}(2) = \sum_{n=n_1-n_2}^{n_1+n_2} \mathfrak{M}_n. \qquad (2n_1+1)(2n_2+1) \qquad (7)$$

Though we shall not prove this result, we can see that it is reasonable. Adding the dimensionalities of the manifolds $\mathfrak{M}_n$ occurring on the right-hand side of (7) we obtain

$$\begin{aligned}[2(n_1-n_2)+1] + [2(n_1-n_2+1)+1] + \ldots + [2(n_1+n_2)+1] \\ = \tfrac{1}{2}\{[2(n_1-n_2)+1] + [2(n_1+n_2)+1]\}(2n_2+1) \\ = (2n_1+1)(2n_2+1),\end{aligned}$$

agreeing, as required, with the dimensionality of $\mathfrak{M}_{n_1}(1) \times \mathfrak{M}_{n_2}(2)$. Secondly, the allowed angular momentum values (6) are plausible by comparison with the corresponding classical result. Classically, angular momentum vectors $\mathbf{M}(1)$ and $\mathbf{M}(2)$ have a resultant $\mathbf{M}$ whose magnitude lies in the range

$$|M(1) - M(2)| \leqslant M \leqslant M(1) + M(2), \qquad (8)$$

as illustrated in *Figure 10*, for $M(1) > M(2)$.

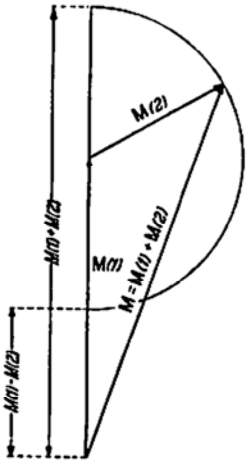

Figure 10. Addition of angular momentum vectors; $M(1) > M(2)$

We shall span the manifold $\mathfrak{M}_n$ by the $(2n+1)$ eigenfunctions of M_z which we shall denote by $\Psi(n_1, n_2; n, m)$, or more briefly by $\Psi(n, m)$, the n_1 and n_2 being understood, and m varying from $-n$ to $+n$. We hence have that

$$\left.\begin{aligned} M^2\Psi(n, m) &= \hbar^2 n(n+1)\Psi(n, m) \\ M_z\Psi(n, m) &= \hbar m\Psi(n, m) \end{aligned}\right\} (m = -n, \ldots n). \qquad (9)$$

Since all functions in $\mathfrak{M}_{n_1}(1) \times \mathfrak{M}_{n_2}(2)$ are linear combinations of the product functions (5), we can, in particular, write the $\Psi(n_1, n_2; n, m)$ in this way:

$$\Psi(n_1, n_2; n, m) = \sum_{m_1=-n_1}^{n_1} \sum_{m_2=-n_2}^{n_2} C^{nm}_{n_1 m_1 n_2 m_2} \Psi_1(n_1, m_1) \Psi_2(n_2, m_2). \tag{10}$$

We can immediately simplify (10). For since $M_z = M_z(1) + M_z(2)$ it follows that

$$M_z \Psi_1(n_1, m_1) \Psi_2(n_2, m_2) = \hbar(m_1 + m_2) \Psi_1(n_1, m_1) \Psi_2(n_2, m_2), \tag{11}$$

i.e. only those terms occur in (10) for which

$$m = m_1 + m_2. \tag{12}$$

We can hence write (10) as

$$\Psi(n_1, n_2; n, m) = \sum_{\substack{m_1+m_2\\=m}} C^{nm}_{n_1 m_1 n_2 m_2} \Psi_1(n_1, m_1) \Psi_2(n_2, m_2). \tag{13}$$

If we fix the normalization and phase factors in $\Psi_k(n_k, m_k)$, then the coefficients $C^{nm}_{n_1 m_1 n_2 m_2}$ are completely determined, apart from an arbitrary phase factor, by the condition that $\Psi(n_1, n_2; n, m)$ should be a normed eigenfunction of M^2 and M_z satisfying (9). These coefficients $C^{nm}_{n_1 m_1 n_2 m_2}$ defining the appropriate linear combinations of wave functions are known as Clebsch–Gordan coefficients or as Wigner coefficients. They can be calculated using the properties of angular momentum operators (see below) or using group theory. But in practice one does not usually have to calculate them as they have been extensively tabulated for the case where the wave functions $\Psi_k(n_k, m_k)$ in (13) are the angular momentum eigenfunctions as defined in section 25, equations (25.22) and (25.25)*. We now see why a consistent choice of phase factors in the wave functions $\Psi_k(n_k, m_k)$ is important, as stated in section 25. We are now taking linear combinations of wave functions, and while an arbitrary phase factor is unimportant in the wave function as a whole, this is of course not the case for individual terms in a wave function. (*Cf* Exercise VI.6, where this difference is illustrated.)

We will now indicate in outline how some of the above results come about. The exact procedure will become clearer from the particular examples which follow.

---

* The Wigner coefficients for $n_2 = \frac{1}{2}, 1, \frac{3}{2}, 2$ are tabulated by Condon and Shortley[21], pp. 76–79; for $n_2 = \frac{5}{2}, 3$ they are tabulated in the *Canadian Journal of Physics*, vol. 30, p. 253 (May 1952). In both these references the Wigner coefficients are written in Dirac notation, as follows

$$C^{nm}_{n_1 m_1 n_2 m_2} = \langle n_1 n_2 m_1 m_2 | n_1 n_2 n m \rangle.$$

van der Waerden[23] gives tables for $n_2 = \frac{1}{2}, 1$. He abbreviates the Wigner coefficient $C^{nm}_{n_1 m_1 n_2 m_2}$ to $C^{n}_{m_1 m_2}$, but the other subscripts and superscripts are of course implied. A notation which has many advantages is $C^{n_1\, n_2\, n}_{m_1 m_2 m}$. Instead of the $C^{n_1\, n_2\, n}_{m_1 m_2 m}$, one often uses $C^{n_1\, n_2 n}_{m_1 m_2}$, defined by $C^{n_1\, n_2\, n}_{m_1 m_2 m} = C^{n_1 n_2 n}_{m_1 m_2}\ \delta_{m, m_1+m_2}$, since the Wigner coefficients vanish unless $m_1 + m_2 = m$.

Equation (11) states that $\Psi_1(n_1, m_1)\Psi_2(n_2, m_2)$ is an eigenfunction of M_z corresponding to the eigenvalue $\hbar m=\hbar(m_1+m_2)$. In general this eigenvalue is degenerate, corresponding to all possible values of m_k, $(m_k=-n_k, \ldots n_k)$, for which $m=m_1+m_2$.

The maximum eigenvalue of M_z is $\hbar(n_1+n_2)$; this eigenvalue is not degenerate; the eigenfunction is $\Psi_1(n_1, n_1)\Psi_2(n_2, n_2)$. As in section 25, the condition

$$A_-A_+\Psi_1(n_1, n_1)\Psi_2(n_2, n_2)=0$$

gives the eigenvalue of M^2 as $\hbar^2(n_1+n_2)(n_1+n_2+1)$. By successively operating on $\Psi_1(n_1, n_1)\Psi_2(n_2, n_2)$ with A_- we generate the $2(n_1+n_2)+1$ eigenfunctions of M_z which span $\mathfrak{M}_{n_1+n_2}$. These eigenfunctions are not yet normalized.

The eigenvalue $\hbar(n_1+n_2-1)$ of M_z is two-fold degenerate; we have two linearly independent eigenfunctions $\Psi_1(n_1, n_1)\Psi_2(n_2, n_2-1)$ and $\Psi_1(n_1, n_1-1)\Psi_2(n_2, n_2)$ and any linear combination of these is an eigenfunction of M_z with eigenvalue $\hbar(n_1+n_2-1)$. One of these linear combinations corresponds to

$$A_-\Psi_1(n_1, n_1)\Psi_2(n_2, n_2) \tag{14}$$

and belongs to $\mathfrak{M}_{n_1+n_2}$ (*i.e.* M has the eigenvalue n_1+n_2). The other linearly independent combination is orthogonal to (14) and corresponds to M having the eigenvalue (n_1+n_2-1), *i.e.* it belongs to $\mathfrak{M}_{n_1+n_2-1}$ and, in fact, is the largest possible eigenvalue of M_z in this manifold. We obtain this linear combination from the condition that

$$A_+\{a\Psi_1(n_1, n_1)\Psi_2(n_2, n_2-1)+b\Psi_1(n_1, n_1-1)\Psi_2(n_2, n_2)\}=0. \tag{15}$$

Starting from this eigenfunction, we can again generate all $2(n_1+n_2-1)+1$ eigenfunctions of M_z spanning $\mathfrak{M}_{n_1+n_2-1}$ by means of the operator A_-. In this way we can continue to generate all eigenfunctions in the product manifold $\mathfrak{M}_{n_1}(1)\times\mathfrak{M}_{n_2}(2)$. In using the operators $A_\pm$ as we have defined them, one must remember that they lead to eigenfunctions which are not yet normed.

In considering some particular examples we introduce the conventional notation: $\mathbf{l}$, $\mathbf{s}$, $\mathbf{j}=\mathbf{l}+\mathbf{s}$, are to denote the orbital, spin and resultant angular momentum vectors of a single particle; $\mathbf{L}$, $\mathbf{S}$, $\mathbf{J}=\mathbf{L}+\mathbf{S}$ are to denote corresponding vectors for a system of particles. We call states in which l has the values 0, 1, 2, 3, . . ., *s*-, *p*-, *d*-, *f*-, *g*-, *h*-, . . . states, and states in which L has these values, *S*-, *P*-, *D*-, *F*-, *G*-, *H*-, . . . states. We specify spin states by the value of $2S+1$, rather than by S itself. $2S+1$ is known as the *multiplicity* of the spin state. For $S=0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ we talk of singlet, doublet, triplet, quartet states, *etc.*

Example 1. The total angular momentum j of an electron with orbital angular momentum l can, according to (6), have the values $l\pm\frac{1}{2}$, which have degeneracy $(2l+2)$ and $2l$ respectively. We considered this case at the end of section 26. On account of the spin-orbit interaction which is different in the two states each single energy level of the spin-free theory splits into two levels, a so-called doublet. In atoms, where the spin-orbit interaction is small (for not too heavy atoms at any rate), these two levels lie very close together. If a magnetic field is applied, these two levels split into $(2l+2)$ and $2l$ levels respectively, corresponding to the different interaction energies of the total magnetic moment of the electron with the magnetic field. In the spin-free theory we should have expected a splitting into $(2l+1)$ levels. s-states ($l=0$) differ from other states, in that, from (6), j can only have the one value $j=\frac{1}{2}$, corresponding to a two-fold degenerate level which splits into a doublet in a magnetic field. The spectra of the alkali metals, which are essentially due to a single electron (the valency electron) in a spherically symmetric field (the Coulomb field of the nucleus, screened by 'closed shells' of electrons), are in agreement with the above interpretation. The two famous D lines of the sodium spectrum represent a typical example. They are due to radiation which the electron emits in falling from an excited p-state ($l=1$) to the ground state, which is an s-state ($l=0$). Whereas the ground state is single ($j=\frac{1}{2}$), there are two excited p-states with different energy levels, corresponding to $j=\frac{3}{2}$ and $j=\frac{1}{2}$. The two observed spectral lines D_1 and D_2 are due to transitions from the $j=\frac{1}{2}$ and $j=\frac{3}{2}$ states respectively.

Example 2. As a second example we consider the resultant spin states of two electrons: $s_1=s_2=\frac{1}{2}$, hence the total spin can have the values $S=0, 1$; a singlet and a triplet state. The former is non-degenerate, the latter three-fold degenerate, giving the four linearly independent spin states which span the 4-dimensional product spin space $\mathfrak{S}_1 \times \mathfrak{S}_2$. Equation (7) we can now write as

$$\mathfrak{S}_1 \times \mathfrak{S}_2 = \mathfrak{M}_1 + \mathfrak{M}_0.$$

We span the three-dimensional manifold $\mathfrak{M}_1$ corresponding to total spin $S=1$, by eigenfunctions of S_z. The eigenfunction corresponding to $S_z=1$ is $\alpha_1\alpha_2$:

$$|11\rangle \qquad \Psi(\tfrac{1}{2}, \tfrac{1}{2}; 1, 1) = \alpha_1\alpha_2. \tag{16a}$$

Since from equations (25.35)

$$A_-(\alpha_1\alpha_2) = \hbar(\alpha_1\beta_2 + \alpha_2\beta_1),$$

it follows that the normed eigenfunction of S_z with eigenvalue 0 is

$$|10\rangle \qquad \Psi(\tfrac{1}{2}, \tfrac{1}{2}; 1, 0) = \frac{1}{\sqrt{2}}(\alpha_1\beta_2 + \alpha_2\beta_1), \tag{16b}$$

and similarly

$$|1{-}1\rangle \qquad \Psi(\tfrac{1}{2}, \tfrac{1}{2}; 1, -1) = \beta_1\beta_2. \tag{16c}$$

We can find the vector defining the one-dimensional manifold $\mathfrak{M}_0$ easily. This vector must be of the form,

$$a\alpha_1\beta_2 + b\alpha_2\beta_1,$$

and it must be orthogonal to $\Psi(\frac{1}{2},\frac{1}{2};1,0)$, equation (16b). This gives at once $a+b=0$, *i.e.*

$|00\rangle$
$$\Psi(\tfrac{1}{2},\tfrac{1}{2};0,0)=\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1). \tag{16d}$$

We could equally well have obtained this vector from one of the conditions

$$A_{\pm}(a\alpha_1\beta_2+b\alpha_2\beta_1)=0,$$

expressing the fact that $\mathfrak{M}_0$ contains no eigenfunctions of S_z with eigenvalues equal to $\pm\hbar$.

We summarize these exceedingly important results: the spin states of two particles of spin $\frac{1}{2}$ divide into a triplet state, $S=1$, with spin wave functions

$$\alpha_1\alpha_2 \tag{16a}$$

$$\frac{1}{\sqrt{2}}(\alpha_1\beta_2+\alpha_2\beta_1) \tag{16b}$$

$$\beta_1\beta_2 \tag{16c}$$

(which are eigenfunctions of S_z belonging to the eigenvalues $+\hbar$, 0, $-\hbar$ respectively), and a singlet state, $S=0$, with spin wave function

$$\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1). \tag{16d}$$

In this second case, the resultant spin is necessarily zero: we can think of the individual spin vectors of the two particles as pointing in opposite directions: we say the spins are antiparallel. We can think of the spins in the triplet states, in which $S=1$, as being parallel. One must of course not interpret this somewhat classical picture too literally. Whereas the triplet spin wave functions are symmetric in the two particles, the singlet spin function is antisymmetric.

Example 3. Finally we consider the combination of the orbital angular momenta of two electrons in *p*-states, *i.e.* $l_1=l_2=1$. The resultant product manifold

$$\mathfrak{M}_1(1)\times\mathfrak{M}_1(2)$$

is spanned by the nine wave functions

$$\Psi_1(1,m_1)\Psi_2(1,m_2) \quad (m_1,m_2=-1,0,1).$$

The resultant orbital angular momentum L assumes, according to (6), the values $L=2, 1, 0$, *i.e.* we have a *D*-state (5-fold degenerate), a *P*-state (3-fold degenerate) and an *S*-state (non-degenerate):

$$\mathfrak{M}_1(1)\times\mathfrak{M}_1(2)=\mathfrak{M}_2+\mathfrak{M}_1+\mathfrak{M}_0.$$

The manifold $\mathfrak{M}_2$ of the $L=2$ states contains the vector

$$\Psi(1,1;2,2)=\Psi_1(1,1)\Psi_2(1,1). \tag{17a}$$

From it, one obtains by repeated application of A_-, using equations (25.23) and (25.24) and normalizing the resultant wave functions,

$$\Psi(1, 1; 2, 1) = \frac{1}{\sqrt{2}}\{\Psi_1(1, 0)\Psi_2(1, 1) + \Psi_1(1, 1)\Psi_2(1, 0)\} \tag{17b}$$

$$\Psi(1, 1; 2, 0) = \frac{1}{\sqrt{6}}\{\Psi_1(1, -1)\Psi_2(1, 1) + 2\Psi_1(1, 0)\Psi_2(1, 0) + \Psi_1(1, 1)\Psi_2(1, -1)\} \tag{17c}$$

$$\Psi(1, 1; 2, -1) = \frac{1}{\sqrt{2}}\{\Psi_1(1, -1)\Psi_2(1, 0) + \Psi_1(1, 0)\Psi_2(1, -1)\} \tag{17d}$$

$$\Psi(1, 1; 2, -2) = \Psi_1(1, -1)\Psi_2(1, -1). \tag{17e}$$

(The normalization of these wave functions is easily obtained, since the individual terms in each wave function are mutually orthogonal. Hence

$$\mathcal{N}\Big\{\sum_{m_1} a_{m_1}\Psi_1(n_1, m_1)\Psi_2(n_2, m - m_1)\Big\} = \sum_{m_1}|a_{m_1}|^2.)$$

The vector $\Psi(1, 1; 1, 1)$ of the manifold $\mathfrak{M}_1$ is orthogonal to $\Psi(1, 1; 2, 1)$, equation (17b), and, furthermore, is the only other vector belonging to the eigenvalue $L_z = 1$. Hence

$$\Psi(1, 1; 1, 1) = \frac{1}{\sqrt{2}}\Big\{\Psi_1(1, 0)\Psi_2(1, 1) - \Psi_1(1, 1)\Psi_2(1, 0)\Big\}, \tag{18a}$$

and applying A_- successively and normalizing, one obtains

$$\Psi(1, 1; 1, 0) = \frac{1}{\sqrt{2}}\Big\{\Psi_1(1, -1)\Psi_2(1, 1) - \Psi_1(1, 1)\Psi_2(1, -1)\Big\} \tag{18b}$$

and

$$\Psi(1, 1; 1, -1) = \frac{1}{\sqrt{2}}\Big\{\Psi_1(1, -1)\Psi_2(1, 0) - \Psi_1(1, 0)\Psi_2(1, -1)\Big\}. \tag{18c}$$

(We could equally well have obtained $\Psi(1, 1; 1, 1)$, equation (18a), by choosing a linear combination such that

$$A_+\{a\Psi_1(1, 1)\Psi_2(1, 0) + b\Psi_1(1, 0)\Psi_2(1, 1)\} = 0$$

which at once gives $a + b = 0$.)

Finally we want the vector $\Psi(1, 1; 0, 0)$ which defines the manifold $\mathfrak{M}_0$. We most easily find this by choosing numbers λ, μ, ν such that either of the equations

$$A_\pm\{\lambda\Psi_1(1, 1)\Psi_2(1, -1) + \mu\Psi_1(1, 0)\Psi_2(1, 0) + \nu\Psi_1(1, -1)\Psi_2(1, 1)\} = 0$$

is satisfied. One at once finds that

$$\lambda + \mu = 0, \qquad \nu + \mu = 0$$

i.e.

$$\Psi(1, 1; 0, 0) = \frac{1}{\sqrt{3}}\Big\{\Psi_1(1, 1)\Psi_2(1, -1) - \Psi_1(1, 0)\Psi_2(1, 0) + \Psi_1(1, -1)\Psi_2(1, 1)\Big\}. \tag{19}$$

Using equations (25.28), (14.4) and (13.26) one easily proves that the wave function $\Psi(1, 1; 0, 0)$ is proportional to the scalar product $\mathbf{r}_1\mathbf{r}_2$ of the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ of the two electrons, and is therefore invariant under rotations of axes. Thus the state $\Psi(1, 1; 0, 0)$, corresponding to zero angular momentum ($L=0$), is spherically symmetric. We had an analogous result for the orbital angular momentum of a single particle, in section 14. In Chapter IX we shall see that this is a quite general result; a system with zero orbital angular momentum has an orbital wave function which is spherically symmetric, *i.e.* it is invariant under rotations of the coordinate axes.

28. Systems of identical particles: Pauli principle

Let us consider a system containing identical particles, say several electrons or protons or neutrons. This means that the two particles are physically indistinguishable. Two electrons, for example, are indistinguishable. We may be able to specify their position as being at two points P_1, and P_2, but it is meaningless to enquire *which* electron is at P_1 and *which* is at P_2. Mathematically, this identity is expressed by the fact that the Hamiltonian of the system is symmetric in the coordinates of the particles: interchanging the coordinates of any two identical particles leaves the Hamiltonian invariant. For simplicity we assume that we deal with two identical particles whose coordinates we label 1 and 2, then

$$H(1, 2) = H(2, 1). \tag{1}$$

We had a typical example of such a symmetric Hamiltonian in Exercise II.2: the Hamiltonian of a helium atom is symmetric in the two electrons.

Let us now consider the eigenvalue problem of the Hamiltonian (1),

$$H(1, 2)\psi(1, 2) = E\psi(1, 2). \tag{2}$$

From (1) it follows (*cf* Exercise VI.10) that if $\psi(1, 2)$ is an eigenfunction belonging to the eigenvalue E, then $\psi(2, 1)$ is also an eigenfunction of the same eigenvalue, *i.e.* E is degenerate and any linear combination

$$\chi(1, 2) = a\psi(1, 2) + b\psi(2, 1) \tag{3}$$

is also an eigenfunction belonging to E.

Not all wave functions (3) represent physically permissible states. Since the two particles are indistinguishable, all physical predictions, such as expectation values and probabilities, must be symmetric in the coordinates of the two particles, *i.e.* they must be independent of which particle we label 1 and which 2. We require, in particular, that for any wave function $\psi(1, 2)$, (3) should imply

$$|\chi(1, 2)|^2 = |\chi(2, 1)|^2. \tag{4}$$

One easily proves (*cf* Exercise VI.11) that (4) implies $a = \pm b$, *i.e.* the allowed wave functions are

$$\chi(1, 2) = \psi(1, 2) \pm \psi(2, 1), \tag{5}$$

i.e. they must be either symmetric or antisymmetric; in the former case, interchanging the particles leaves the wave function invariant, in the latter case, interchanging the particles multiplies the wave function by (-1). We can express this more elegantly if we introduce the particle *exchange* or *permutation* operator P_{12} defined by

$$P_{12}\psi(1, 2) = \psi(2, 1). \tag{6}$$

Symmetric and antisymmetric wave functions then have the property that

$$P_{12}\chi(1, 2) = \pm\chi(1, 2) \tag{7}$$

respectively, *i.e.* they are eigenfunctions of P_{12} with the eigenvalues ± 1 respectively.

It follows from the time-dependent Schrödinger equation that the symmetry character of a wave function persists throughout time. For

$$i\hbar\frac{\partial}{\partial t}\psi_t(1, 2) = H(1, 2)\psi_t(1, 2)$$

implies that

$$P_{12}d\psi_t(1, 2) = \frac{dt}{i\hbar}H(1, 2)\psi_t(2, 1) = \pm\frac{dt}{i\hbar}H(1, 2)\psi_t(1, 2) = \pm d\psi_t(1, 2),$$

according as $P_{12}\psi_t(1, 2) = \pm\psi_t(1, 2)$, *i.e.* the wave function remains symmetric or antisymmetric throughout time: its symmetry character is a constant of the motion. We can also see this directly from the operators:

$$[P_{12}, H(1, 2)] = 0, \tag{8}$$

since for any wave function $\psi(1, 2)$

$$P_{12}H(1, 2)\psi(1, 2) = H(2, 1)\psi(2, 1) = H(1, 2)P_{12}\psi(1, 2).$$

The operator P_{12} exchanges the particles 1 and 2, *i.e.* it exchanges all their coordinates, both spatial and spin. It is often convenient to use position and spin exchange operators $P_{12}{}^x$ and $P_{12}{}^\sigma$ which respectively exchange the position coordinates and spin coordinates only, and clearly

$$P_{12} = P_{12}{}^x P_{12}{}^\sigma = P_{12}{}^\sigma P_{12}{}^x. \tag{9}$$

Thus if $\Phi(\mathbf{r}_1, s_z^{(1)}; \mathbf{r}_2, s_z^{(2)})$ denotes a two-electron wave function then

$$P_{12}{}^x\Phi(\mathbf{r}_1, s_z^{(1)}; \mathbf{r}_2, s_z^{(2)}) = \Phi(\mathbf{r}_2, s_z^{(1)}; \mathbf{r}_1, s_z^{(2)}),$$

$$P_{12}{}^\sigma\Phi(\mathbf{r}_1, s_z^{(1)}; \mathbf{r}_2, s_z^{(2)}) = \Phi(\mathbf{r}_1, s_z^{(2)}; \mathbf{r}_2, s_z^{(1)}),$$

and

$$P_{12}\Phi(\mathbf{r}_1, s_z^{(1)}; \mathbf{r}_2, s_z^{(2)}) = \Phi(\mathbf{r}_2, s_z^{(2)}; \mathbf{r}_1, s_z^{(1)}).$$

So far we have shown that wave functions of identical particles are either symmetric or antisymmetric. In practice one finds that the wave functions for a particular kind of elementary particle are either always symmetric or always antisymmetric, *i.e.* only if we restrict the wave functions in this way do we obtain agreement with observation. Thus the wave functions of particles of spin $\frac{1}{2}$ such as electrons, positrons, protons and neutrons are always antisymmetric and the wave functions of particles with spin zero (π-mesons) or with spin 1 (photons) are always symmetric. This distinction between particles whose wave functions are symmetric and antisymmetric is most fundamental and leads to different behaviour if we consider an assembly of particles, *i.e.* it leads to different statistics. Particles with integral spin (symmetric wave functions) satisfy Bose–Einstein statistics, particles with 'half odd-integral' spin (antisymmetric wave functions) satisfy Fermi–Dirac statistics. Consequently they are often called bosons and fermions respectively.

The Hamiltonian of two non-interacting identical particles separates into two parts

$$H(1, 2) = H(1) + H(2).$$

If

$$\varphi_1(k), \varphi_2(k), \ldots, \varphi_n(k), \ldots, \quad (k = 1, 2)$$

are the eigenfunctions of $H(k)$, corresponding to eigenvalues

$$E_1, E_2, \ldots E_m, \ldots,$$

then the eigenfunction of $H(1, 2)$, belonging to the eigenvalue $E_m + E_n$ is

$$\Phi_{mn}(1, 2) = \frac{1}{\sqrt{2}}\left\{\varphi_m(1)\varphi_n(2) \pm \varphi_m(2)\varphi_n(1)\right\}. \qquad (10)$$

If the particles are fermions, we must take the minus sign in (10). Hence for $m = n$, $\Phi_{mn}(1, 2) \equiv 0$: *it is not possible for two identical fermions to be in identical states.* This is the original formulation of *Pauli's exclusion principle.* However, if the particles interact, it is no longer possible to separate the Hamiltonian and the unsymmetrized eigenfunctions are no longer products of one-particle eigenfunctions. Hence we can then no longer formulate the Pauli principle in this way. But our original statement that the wave function must be antisymmetric is a generally valid formulation of Pauli's exclusion principle. Nevertheless, even in the case of interacting fermions, it is often useful to talk of individual particle states and product wave functions as a first approximation, provided the interactions are small perturbations. The theory of atomic structure represents an excellent example where this approximation is applicable, as the Coulomb repulsion between the atomic electrons is usually small compared to the interaction of the atomic electrons with the nucleus.

non-identity

We have so far mainly considered two identical particles, but the ideas we have developed can easily be generalized to any number of particles.

In section 27 we derived the wave functions of systems containing several particles with definite angular momentum properties. We now illustrate, for some particular atomic and nuclear systems, how this procedure is to be modified if the particles satisfy the Pauli exclusion principle.

To begin with, we remind the reader that the triplet spin wave functions of two particles of spin $\frac{1}{2}$ are symmetric in the spin coordinates, whereas the singlet spin wave function is antisymmetric [equations (27.16 a–d)]:

$$P_{12}\alpha_1\alpha_2=\alpha_1\alpha_2,\ etc,\quad P_{12}\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1)=-\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1).$$

Example 1. As first example we consider two atomic electrons, the spatial wave functions of which are $\varphi_a(1)$ and $\varphi_b(2)$, so that the spatial wave function of the combined system, apart from its symmetry properties, is $\varphi_a(1)\varphi_b(2)$. Combining this wave function with the spin wave functions (27.16a–d) we can form four antisymmetric combinations. We obtain the singlet state ($S=0$)

$$\frac{1}{\sqrt{2}}[\varphi_a(1)\varphi_b(2)+\varphi_a(2)\varphi_b(1)]\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1) \tag{11a}$$

and the three triplet states ($S=1$)

$$\frac{1}{\sqrt{2}}[\varphi_a(1)\varphi_b(2)-\varphi_a(2)\varphi_b(1)]\times\begin{cases}\alpha_1\alpha_2 & \text{(11b)}\\ \frac{1}{\sqrt{2}}(\alpha_1\beta_2+\alpha_2\beta_1) & \text{(11c)}\\ \beta_1\beta_2\quad , & \text{(11d)}\end{cases}$$

and these wave functions are normalized if φ_a and φ_b are normalized.

If the two electrons are in the same s-state (we call this an s^2-configuration), then $\varphi_a(r)\equiv\varphi_b(r)$ and only the singlet state (11a) is possible. We denote this state by the symbol ${}^1S_0$, the general notation being ${}^{2S+1}L_J$ where $2S+1$ is the spin multiplicity of the state and J its resultant angular momentum.

If we consider a p^2-configuration we have all together 36 linearly independent product wave functions (three spatial and two spin, giving six total wave functions for each particle, and the total wave functions for the two particles can be combined in any order). We easily enumerate the antisymmetric states. Firstly, we see that there are 15 such states only: the total wave function of the first electron can be chosen in 6 ways, that of the second in only 5 ways; this gives 30 possibilities; but the order of the electrons is immaterial since we antisymmetrize the wave function: hence there are 15 possible states. To classify these we remember that triplet spin wave functions are symmetric, singlet spin wave functions antisymmetric, and that the p^2-configuration gives states with $L=2, 1, 0$. Of these the $L=2$ and $L=0$ states are symmetric, the $L=1$ state antisymmetric (*cf* section 27, example 3, particularly equations (27.17–19)). Combining these results we find the possible antisymmetric states by combining symmetric spin states with antisymmetric spatial wave functions, and vice versa. In this way we obtain the following states:

$$ {}^1D_2,\ {}^1S_0,\ \ {}^3P_0,\ {}^3P_1,\ {}^3P_2. \tag{12}$$

We obtain three different P states, as the spin $S=1$ can combine with the orbital angular momentum $L=1$ to give resultants $J=0$, 1 or 2. One easily checks that the total number of states in (12) is 15 (${}^{2S+1}L_J$ corresponds of course to $2J+1$ states). It is not difficult to obtain the explicit wave functions of these states, using equations (27.17–19) and (27.16), but we shall not write down these rather lengthy formulae. The reader should have no difficulty in deriving them if he so desires. Methods of finding the antisymmetric configurations of more complicated systems have been given by van der Waerden[23].

Example 2. As second example we consider the ground state wave function of the helium nucleus or α-particle. The α-particle consists of two neutrons and two protons and the ground state is a ${}^1S_0$ state, if one assumes the forces between the nucleons (protons and neutrons being the basic constituents of nuclei are called nucleons) to be central forces.

Let 1 and 2 denote the protons and 3 and 4 the neutrons. Combining the proton spin states we get a triplet and a singlet state, and similarly for the neutron spin states. To obtain a singlet state ($S=0$) for the α-particle we must then either combine the two singlet states giving the spin wave function

$$\xi_a(\widetilde{12};\widetilde{34}) = \tfrac{1}{2}(\alpha_1\beta_2-\alpha_2\beta_1)(\alpha_3\beta_4-\alpha_4\beta_3), \tag{13}$$

or we must combine the triplet spin wave functions analogously to the way we obtained an $L=0$ state from a p^2-configuration in equation (27.19). The resulting spin function is

$$\xi_s(\overline{12};\overline{34}) = \frac{1}{\sqrt{3}}\Big\{\alpha_1\alpha_2\beta_3\beta_4 - \tfrac{1}{2}(\alpha_1\beta_2+\alpha_2\beta_1)(\alpha_3\beta_4+\alpha_4\beta_3) + \beta_1\beta_2\alpha_3\alpha_4\Big\}. \tag{14}$$

The spin states (13) and (14) are respectively antisymmetric and symmetric in the like particles (we indicated this by the suffixes a and s, and by placing tildes and bars over pairs of particles in which the wave functions are antisymmetric and symmetric respectively). Hence the total wave function of the α-particle is of the form

$$\Phi(\widetilde{12};\widetilde{34}) = a\varphi_s(\overline{12};\overline{34})\xi_a(\widetilde{12};\widetilde{34}) + b\varphi_a(\widetilde{12};\widetilde{34})\xi_s(\overline{12};\overline{34}) \tag{15}$$

where φ_s and φ_a are spatial wave functions with the indicated symmetry properties and a and b are constants specifying the mixture of symmetric and antisymmetric spin states. By comparing the calculated with the experimentally observed binding energies of the ground state, one finds that $|b| \ll |a|$, *i.e.* in the ground state the protons and neutrons are predominantly paired off by themselves into singlet states.

EXERCISES VI

VI.1. Prove equations (25.12) to (25.14).

VI.2. Prove equation (25.24).

VI.3. What is the component of the Pauli spin vector $\boldsymbol{\sigma}$ in the direction of the unit vector $\mathbf{n} = (n_1, n_2, n_3)$?

VI.4. Denoting the Pauli matrices by σ_i, $i = 1, 2, 3$, instead of $\sigma_x, \sigma_y, \sigma_z$, prove the anticommutation relations

$$\sigma_i\sigma_k + \sigma_k\sigma_i = 2\delta_{ik}.$$

VI.5. Prove that any 2×2 Hermitian matrix can be written as a linear combination of the 2×2 unit matrix and the Pauli matrices.

VI.6. (*i*) If σ_n is the component of $\boldsymbol{\sigma}$ in the direction of the unit vector $\mathbf{n}$ with components $(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, find the eigenfunctions of σ_n referred to the vectors α and β as basis $(\sigma_z\alpha = \alpha, \sigma_z\beta = -\beta, (\alpha, \alpha) = 1, (\beta, \beta) = 1)$.

(*ii*) If $\Psi = \psi_1(\mathbf{r})\alpha + \psi_2(\mathbf{r})\beta$ is a normed wave function, what are the probabilities of observing a component of spin $\pm\hbar/2$ in the directions $\pm\mathbf{n}$ respectively?

(*iii*) If we are not given the complete wave function Ψ, but only know that the system is in a state in which s_z has the values $\pm\hbar/2$ with probabilities $|c_1|^2$ and $|c_2|^2$, $(|c_1|^2 + |c_2|^2 = 1)$, what are the probabilities of observing a component of spin $\pm\hbar/2$ in the directions $\pm\mathbf{n}$ respectively?

VI.7. If $\mathbf{M} = \mathbf{M}(1) + \mathbf{M}(2)$, and $\mathbf{M}(k)$, $(k = 1, 2)$ satisfy the commutation relations (25.6–7) and (25.12–14), prove that $\mathbf{M}$ also satisfies these relations.

VI.8. Obtain the total angular momentum eigenfunctions of the operators j^2 and j_z of a single particle of spin $\frac{1}{2}$ in a *p*-state.

VI.9. Given three particles of spin $\frac{1}{2}$, $s_1 = s_2 = s_3 = \frac{1}{2}$, obtain the spin wave functions of the combined system which are eigenfunctions of S and S_z.

VI.10. Prove that if $H(1, 2) = H(2, 1)$ and $H(1, 2)\psi(1, 2) = E\psi(1, 2)$, then $\psi(2, 1)$ is an eigenfunction of $H(1, 2)$ belonging to the same eigenvalue E.

VI.11. Prove equation (28.5) from the indistinguishability condition (28.4).

VI.12. The nucleus of the 'helium three' isotope, written ^{3}He, consists of two protons and one neutron. Its ground state, assuming central forces, is a $^2S_{\frac{1}{2}}$ state. Write the ground state wave function in a form analogous to the α-particle wave function (28.15).

VII

TIME-INDEPENDENT PERTURBATION THEORY

THERE exist several approximate methods of solution for eigenvalue problems which are too complicated to admit exact solutions. In this chapter, we shall consider one of the most important of these, perturbation theory. More particularly, we shall consider the application of perturbation theory to the calculation of the stationary energy levels and the corresponding energy eigenfunctions of systems (such as atoms or molecules) which are too complicated to permit a rigorous solution.

29. First-order perturbation: non-degenerate case

The basic idea underlying the application of perturbation theory is the following. Corresponding to the operator H, whose eigenvalue problem we want to solve, we suppose that there exists an operator $H^{(0)}$ differing by only a small term from H, and that the complete solution of the eigenvalue problem of $H^{(0)}$ is known. We write

$$H = H^{(0)} + \varepsilon V \tag{1}$$

and, for definiteness, we shall assume that V is of the same order of magnitude as $H^{(0)}$ and that ε is small compared to unity. The state of affairs here depicted exists in not too heavy atoms; the Coulomb repulsion between the electrons represents a small perturbation compared to the interaction of the electrons with the nucleus.

Suppose

$$\left.\begin{array}{l} \varphi_1{}^{(0)}, \varphi_2{}^{(0)}, \ldots \varphi_r{}^{(0)}, \ldots \\ E_1{}^{(0)}, E_2{}^{(0)}, \ldots E_r{}^{(0)}, \ldots \end{array}\right\} \tag{2}$$

are the complete sets of eigenfunctions and eigenvalues of the unperturbed eigenvalue problem,

$$H^{(0)} \varphi^{(0)} = E^{(0)} \varphi^{(0)}, \tag{3}$$

which are assumed known. On the basis of continuity arguments, it is then reasonable to assume that corresponding to each eigensolution $\varphi_r{}^{(0)}$ and eigenvalue $E_r{}^{(0)}$ of the unperturbed problem, there exists an eigenfunction φ_r and eigenvalue E_r of the perturbed (exact) eigenvalue problem,

$$H\varphi = E\varphi \tag{4}$$

which are nearly equal to the unperturbed quantities, differing only by quantities of the order of magnitude ε, and the resulting set of eigenfunctions $\varphi_1, \varphi_2, \ldots \varphi_n, \ldots$ is also complete[*]. Corresponding to a particular unperturbed eigenfunction $\varphi_N^{(0)}$ and eigenvalue $E_N^{(0)}$ we can then try to express φ_N and E_N as power series in ε:

$$\varphi_N = \varphi_N^{(0)} + \varepsilon\varphi_N^{(1)} + \varepsilon^2\varphi_N^{(2)} + \ldots \tag{5}$$

$$E_N = E_N^{(0)} + \varepsilon E_N^{(1)} + \varepsilon^2 E_N^{(2)} + \ldots \tag{6}$$

where $\varepsilon^n\varphi_N^{(n)}$ and $\varepsilon^n E_N^{(n)}$ are the nth order corrections to the unperturbed wave functions and eigenvalues.

For the time being, we assume that all the unperturbed eigenvalues $E_1^{(0)}, E_2^{(0)}, \ldots E_r^{(0)}, \ldots$ are non-degenerate, so that the corresponding eigenfunctions, assumed normed, are completely determined apart from unimportant phase-factors.

Substituting equations (5) and (6) into (4) we obtain

$$(H^{(0)} + \varepsilon V)(\varphi_N^{(0)} + \varepsilon\varphi_N^{(1)} + \varepsilon^2\varphi_N^{(2)} + \ldots)$$
$$= (E_N^{(0)} + \varepsilon E_N^{(1)} + \ldots)(\varphi_N^{(0)} + \varepsilon\varphi_N^{(1)} + \ldots). \tag{7}$$

Equating coefficients of powers of ε, we obtain the following set of equations from the terms in ε^0, ε^1 and ε^2:

0 order

$$H^{(0)}\varphi_N^{(0)} = E_N^{(0)}\varphi_N^{(0)}, \tag{8a}$$

1st order

$$H^{(0)}\varphi_N^{(1)} + V\varphi_N^{(0)} = E_N^{(0)}\varphi_N^{(1)} + E_N^{(1)}\varphi_N^{(0)} \tag{8b}$$

2nd order

$$H^{(0)}\varphi_N^{(2)} + V\varphi_N^{(1)} = E_N^{(0)}\varphi_N^{(2)} + E_N^{(1)}\varphi_N^{(1)} + E_N^{(2)}\varphi_N^{(0)}. \tag{8c}$$

Equation (8a) is identically satisfied, as it must be, since it represents the zero-order approximation of our perturbation problem.

To obtain the first-order approximation we must take into account equation (8b) but neglect equation (8c). We must determine the first-order corrections $\varphi_N^{(1)}$ and $E_N^{(1)}$ to the wave function and the energy so that equation (8b) holds. For this purpose we expand $\varphi_N^{(1)}$ in terms of the unperturbed eigenfunctions $\varphi_r^{(0)}$,

$$\varphi_N^{(1)} = \sum_{r=1}^{\infty} a_r\varphi_r^{(0)}. \tag{9}$$

Substituting this expansion into equation (8b), we obtain, using equation (3),

$$\sum_{r=1}^{\infty} a_r E_r^{(0)}\varphi_r^{(0)} + V\varphi_N^{(0)} = E_N^{(0)}\sum_{r=1}^{\infty} a_r\varphi_r^{(0)} + E_N^{(1)}\varphi_N^{(0)}. \tag{10}$$

Taking the inner product with $\varphi_s^{(0)}$, equation (10) gives

$$a_s E_s^{(0)} + (\varphi_s^{(0)}, V\varphi_N^{(0)}) = a_s E_N^{(0)} + E_N^{(1)}\delta_{sN}, \quad (s = 1, 2, \ldots). \tag{11}$$

---

[*] For simplicity we have adapted our notation to the case where both $H^{(0)}$ and H have completely discrete spectra of eigenvalues. But the following arguments also cover the case of a partly continuous spectrum, by replacing summations over eigenvalues by integrations over appropriate ranges of eigenvalues. If, as in the present chapter, we are interested in the shift due to the perturbation of a given energy level $E_r^{(0)}$, it is of course implicit that this level belongs to the discrete spectrum.

Abbreviating the matrix elements of the perturbation V by

$$V_{rs} = (\varphi_r^{(0)}, V\varphi_s^{(0)}), \tag{12}$$

equations (11) give for the cases $s = \mathcal{N}$ and $s \neq \mathcal{N}$:

$$E_N^{(1)} = V_{NN} \tag{13}$$

and

$$a_s = \frac{V_{sN}}{E_N^{(0)} - E_s^{(0)}}, \quad (s \neq \mathcal{N}). \tag{14}$$

Equation (13) gives the first-order correction to the energy level. It is seen to be the expectation value of the perturbation operator in the unperturbed state $\varphi_N^{(0)}$ and thus depends only on the zero-order wave functions. Equations (14) give the expansion coefficients a_s of the first-order correction to the wave function, equation (9), *except for the coefficient* a_N of the unperturbed wave function $\varphi_N^{(0)}$ of the state considered. This coefficient is determined from the normalization of the wave function (5) to terms of order ε:

$$\begin{aligned}(\varphi_N, \varphi_N) &= (\varphi_N^{(0)} + \varepsilon\varphi_N^{(1)} + \ldots, \varphi_N^{(0)} + \varepsilon\varphi_N^{(1)} + \ldots) \\ &= (\varphi_N^{(0)}, \varphi_N^{(0)}) + \varepsilon[(\varphi_N^{(1)}, \varphi_N^{(0)}) + (\varphi_N^{(0)}, \varphi_N^{(1)})] \\ &\qquad + (\text{terms in } \varepsilon^2, \varepsilon^3, \ldots).\end{aligned} \tag{15}$$

Since $(\varphi_N^{(0)}, \varphi_N^{(0)}) = 1$, it follows from (15) and $a_N = (\varphi_N^{(0)}, \varphi_N^{(1)})$ that we require

$$a_N + \bar{a}_N = 0, \tag{16}$$

i.e. a_N is a purely imaginary number which we shall write

$$a_N = i\gamma_N \quad (\gamma_N \text{ real}). \tag{17}$$

We can hence write the eigenfunction φ_N, correct to the first order in ε, in the form

$$\varphi_N = \varphi_N^{(0)}\ (1 + i\varepsilon\gamma_N) + \varepsilon\sum_r{}' \frac{V_{rN}}{E_N^{(0)} - E_r^{(0)}}\varphi_r^{(0)}, \tag{18}$$

where $\sum_r{}'$ denotes summation over all values of r, except $r = \mathcal{N}$ Since to terms of the order of ε, $\exp(i\gamma_N\varepsilon) = 1 + i\gamma_N\varepsilon$, we can also write the first term in (18) as

$$\exp(i\gamma_N\varepsilon)\varphi_N^{(0)}.$$

We now see that the choice of γ_N in (17), which we have so far left open, simply corresponds to a particular choice of phase factor of the perturbed wave function φ_N and as such is unimportant. We can, without loss of generality, put

$$a_N = i\gamma_N = 0, \tag{19}$$

and write the first-order wave function (18) as

$$\varphi_N = \varphi_N^{(0)} + \varepsilon\sum_r{}' \frac{V_{rN}}{E_N^{(0)} - E_r^{(0)}}\varphi_r^{(0)}. \tag{20}$$

30. First-order perturbation: degenerate case

We must now consider the case, which frequently arises in practice, that the unperturbed energy eigenvalue $E_N^{(0)}$, in which we are interested, is degenerate. The corresponding manifold of eigenvectors of $H^{(0)}$ belonging to the eigenvalue $E_N^{(0)}$ can then be spanned in infinitely many ways and it is not apparent which linear combinations are the correct zero-order wave functions.

We can see in yet another way that the treatment of section 29 breaks down if $E_N^{(0)}$ is degenerate. For if $E_r^{(0)} = E_N^{(0)}$, for some $r \neq N$, then the corresponding denominators $(E_N^{(0)} - E_r^{(0)})$ in equation (29.20) vanish, so that this equation becomes meaningless. But we can see from this equation how this difficulty is to be overcome. We must choose the zero order eigenfunctions corresponding to the eigenvalue $E_N^{(0)}$ so that

$$V_{rN} = 0 \quad (r \neq N) \tag{1}$$

whenever $E_r^{(0)} = E_N^{(0)}$. The troublesome terms in (29.20) then vanish identically.

Suppose the eigenvalue $E_N^{(0)}$ is $(a+1)$-fold degenerate:

$$E_N^{(0)} = E^{(0)}_{N+1} = \dots = E^{(0)}_{N+a}. \tag{2}$$

The corresponding eigenfunctions of $H^{(0)}$ define an $(a+1)$-dimensional manifold $\mathfrak{M}^{(0)}$. We want to find a set of $(a+1)$ orthonormal functions $\varphi_N^{(0)}, \varphi^{(0)}_{N+1}, \dots \varphi^{(0)}_{N+a}$, spanning $\mathfrak{M}^{(0)}$ such that (1) holds for them:

$$V_{N+j,\, N+k} = (\varphi^{(0)}_{N+j}, V\varphi^{(0)}_{N+k}) = 0, \quad (j \neq k;\, j, k = 0, 1, \dots a), \tag{3}$$

i.e. referred to $\varphi_N^{(0)}, \varphi^{(0)}_{N+1}, \dots \varphi^{(0)}_{N+a}$ as basis in $\mathfrak{M}^{(0)}$, $V_{N+j,\, N+k}$ is a diagonal matrix,

$$V_{N+j,\, N+k} = (\varphi^{(0)}_{N+j}, V\varphi^{(0)}_{N+k}) = v_{jk} = v_k\, \delta_{jk}, \quad (j, k = 0, 1, \dots a), \tag{4}$$

where

$$v_k = (\varphi^{(0)}_{N+k}, V\varphi^{(0)}_{N+k}), \quad (k = 0, 1, \dots a). \tag{5}$$

Suppose we initially choose an arbitrary set of $(a+1)$ orthonormal eigenfunctions $\xi_0, \xi_1, \dots \xi_a$, to span $\mathfrak{M}^{(0)}$. Our problem is then to find a unitary transformation U in $\mathfrak{M}^{(0)}$, *i.e.* an $(a+1) \times (a+1)$ unitary matrix, which transforms the basis $\xi_0, \xi_1, \dots \xi_a$ into the correct linear combinations $\varphi_N^{(0)}, \varphi^{(0)}_{N+1}, \dots \varphi^{(0)}_{N+a}$ to act as new basis in $\mathfrak{M}^{(0)}$. Apart from some slight changes in notation, this is exactly the eigenvalue problem which we solved in section 22 (*d*): to diagonalize the operator V in the manifold $\mathfrak{M}^{(0)}$. Unlike the case in section 22, $\mathfrak{M}^{(0)}$ now only forms a part of our function space. As a result, v_k and $\varphi^{(0)}_{N+k}$, $(k = 0, 1, \dots a)$, are eigenvalues and eigenfunctions of V in $\mathfrak{M}^{(0)}$ *only*, but not, in general, of the *whole* function space. This follows from equation (4), which implies that

$$V\varphi^{(0)}_{N+k} = v_k\, \varphi^{(0)}_{N+k} + \chi_{N+k} \quad (k = 0, 1, \dots a), \tag{6}$$

where χ_{N+k} denotes a vector orthogonal to the whole manifold $\mathfrak{M}^{(0)}$.

$$\begin{vmatrix} v_j - E' & v_{jk} \\ v_{kj} & v_k - E' \end{vmatrix} = 0$$

If we have solved the eigenvalue problem of V in $\mathfrak{M}^{(0)}$ by the methods of section 22 (d), we take the above linear combinations $\varphi^{(0)}_{N+k}$, $(k=0, 1, \ldots a)$ as zero-order wave functions corresponding to the eigenvalue $E_N{}^{(0)}$ and proceed exactly as in section 29. Writing the first-order eigenfunctions and eigenvalues of H corresponding to the zero-order energy level $E_N{}^{(0)}$ as

$$\varphi_{N+k}=\varphi^{(0)}_{N+k}+\varepsilon\varphi^{(1)}_{N+k}, \quad (k=0, 1, \ldots a), \tag{7}$$

$$E_{N+k}=E_N{}^{(0)}+\varepsilon E^{(1)}_{N+k}, \quad (k=0, 1, \ldots a), \tag{8}$$

we obtain, analogously to equations (29.13) and (29.20),

$$E^{(1)}_{N+k}=V_{N+k,\,N+k}=v_k, \quad (k=0, 1, \ldots a), \tag{9}$$

$$\varphi_{N+k}=\varphi^{(0)}_{N+k}+\varepsilon\sum_r{}' \frac{V_{r,N+k}}{E^{(0)}_{N+k}-E_r{}^{(0)}}\varphi_r{}^{(0)}, \quad (k=0, 1, \ldots a), \tag{10}$$

where $\sum_r{}'$ now denotes summation over all values of r except $r=N$, $N+1, \ldots N+a$.

Equation (9) gives the first-order corrections to the energy eigenvalues,

$$E_{N+k}=E_N{}^{(0)}+\varepsilon v_k, \quad (k=0, 1, \ldots a). \tag{11}$$

The first-order corrections again only depend on the zero-order wave functions and not on the first-order corrections to these. In general, the eigenvalues $v_0, v_1, \ldots v_a$ of V in $\mathfrak{M}^{(0)}$ will not all also be degenerate, so that some (even if not all) of the v_k will be different from each other. As a result, the $(a+1)$-fold degenerate zero-order energy level $E_N{}^{(0)}$ is, in general, split into several [at most $(a+1)$] distinct levels, whose positions are given by equation (11). Equation (10) gives the corresponding first-order eigenfunctions. One easily verifies that, as in the non-degenerate case, these wave functions are correctly normalized.

In Chapter IX we shall return to the determination of the correct zero-order wave functions for perturbation calculations of degenerate levels and splitting of these levels by perturbations. We shall find that these problems can be treated using general qualitative arguments depending on the symmetry properties of the unperturbed Hamiltonian $H^{(0)}$ and the perturbation εV. Furthermore, many of these results will be true rigorously, unlike the approximate results of this chapter.

The first-order perturbation theory which we have developed in the last two sections represents those parts of the theory which are most frequently required in practice. One can derive expressions for the higher order perturbations but these become increasingly complicated and are therefore often of little use in practice. We shall consider second-order perturbation theory in Exercise VII.4.

31. Simple examples on perturbation theory

We shall now illustrate by some examples the perturbation theory which has been developed. We shall only consider very simple examples, as the practical problems usually lead to very heavy analysis and computation.

Example 1 : *Ground state of the helium atom*

In Chapter III (Exercise III.6) we found the wave function of the ground state of helium to be

$$\varphi^{(0)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{8a^3}{\pi} e^{-2a(r_1+r_2)}, \tag{1}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of the two electrons relative to the nucleus, provided the nucleus is taken as infinitely heavy (*i.e.* we neglect its motion) and provided the Coulomb repulsion

$$V = \frac{e^2}{r_{12}} \quad (r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|) \tag{2}$$

between the electrons is neglected. The corresponding energy of the ground state is

$$E^{(0)} = -4e^2 a, \tag{3}$$

$a^{-1} = \hbar^2/me^2$ being the Bohr radius of the hydrogen atom. We now want to calculate the correction to the energy level (3) of the ground state due to the Coulomb interaction (2), using first-order perturbation theory. From section 29, the corrected energy eigenvalue is

$$E = E^{(0)} + (\varphi^{(0)}, V\varphi^{(0)}) \tag{4}$$

where

$$(\varphi^{(0)}, V\varphi^{(0)}) = \frac{64a^6}{\pi^2} \int \frac{e^2}{r_{12}} e^{-4a(r_1+r_2)} d\mathbf{r}_1 d\mathbf{r}_2 = \frac{5}{4} e^2 a. \tag{5}$$

(This six-fold integral has been evaluated by Pauling and Wilson[11] Appendix V.) Hence first-order perturbation theory gives the energy of the ground state as

$$E = -2{\cdot}75 e^2 a.$$

This represents a considerable improvement on the zero-order approximation, equation (3), since the experimental value is $-2{\cdot}90 e^2 a$.

We have been able to neglect spin completely in this analysis, since we have assumed the unperturbed Hamiltonian and the perturbation spin-independent. The antisymmetric zero-order wave function, including spin, is

$$\varphi^{(0)}(\mathbf{r}_1, \mathbf{r}_2) \frac{1}{\sqrt{2}} (\alpha_1 \beta_2 - \alpha_2 \beta_1), \tag{6}$$

but the spin function does not alter the value of the matrix element (5).

Example 2: Two identical particles

We consider the more general case of two identical particles (position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$), described by the Schrödinger equation

$$(H^{(0)}+V)\varphi(\mathbf{r}_1,\mathbf{r}_2)=E\varphi(\mathbf{r}_1,\mathbf{r}_2), \tag{7}$$

where

$$V=V(\mathbf{r}_1,\mathbf{r}_2)=V(\mathbf{r}_2,\mathbf{r}_1) \tag{8}$$

is the interaction between the particles which we assume to depend on $\mathbf{r}_1$ and $\mathbf{r}_2$ only; on account of the identity of the particles it must be symmetric in $\mathbf{r}_1$ and $\mathbf{r}_2$. We shall treat V as a perturbation of the unperturbed Hamilton $H^{(0)}$ which is separable,

$$H^{(0)}=H(1)+H(2) \tag{9}$$

with

$$H(k)=H(\mathbf{r}_k,\ -i\hbar\ \text{grad}_k),\quad (k=1,2), \tag{10}$$

describing the motion of the kth particle in an external field (*e.g.* two electrons in an atom; in this case V is their mutual Coulomb repulsion).

If $u_n(k)$, ε_n, $(n=1,2,3,\ldots)$, are the eigenfunctions and eigenvalues of $H(k)$, *i.e.*

$$H(k)u_n(k)=\varepsilon_n u_n(k),\quad \begin{pmatrix}k=1,2\\ n=1,2,3,\ldots\end{pmatrix}, \tag{11}$$

then

$$E^{(0)}=\varepsilon_m+\varepsilon_n,\quad (m\neq n,\ m,n=1,2,3,\ldots) \tag{12}$$

is a two-fold degenerate eigenvalue of the unperturbed Hamiltonian (9), with eigenfunctions

$$\xi_1(\mathbf{r}_1,\mathbf{r}_2)=u_m(1)u_n(2),\quad \xi_2(\mathbf{r}_1,\mathbf{r}_2)=u_m(2)u_n(1). \tag{13}$$

To calculate the effect of the interaction V on the level $E^{(0)}$, in first-order perturbation, we must diagonalize the matrix

$$[V_{ij}]=(\xi_i,V\xi_j),\quad (i,j=1,2), \tag{14}$$

in the two-dimensional manifold spanned by ξ_1 and ξ_2. On account of the symmetry of V, equation (8), one easily shows that $V_{11}=V_{22}$ and $V_{12}=V_{21}$, so that the matrix $[V_{ij}]$ has the particularly symmetric form

$$[V_{ij}]=\begin{bmatrix}V_{11} & V_{12}\\ V_{12} & V_{11}\end{bmatrix}. \tag{15}$$

The diagonalization of this matrix is exactly the problem treated in Exercise V.5. The eigenfunctions, the correct zero-order wave functions for a perturbation calculation, are the symmetric and antisymmetric combinations

$$\varphi_1{}^{(0)}=\frac{1}{\sqrt{2}}\{u_m(1)u_n(2)+u_m(2)u_n(1)\} \tag{16a}$$

$$\varphi_2{}^{(0)}=\frac{1}{\sqrt{2}}\{u_m(1)u_n(2)-u_m(2)u_n(1)\} \tag{16b}$$

and the corresponding energy levels are given by

$$E_i = E^{(0)} + E_i^{(1)} = (\varepsilon_m + \varepsilon_n) + E_i^{(1)}, \quad (i = 1, 2) \tag{17}$$

with

$$E_i^{(1)} = (\varphi_i^{(0)}, V\varphi_i^{(0)}) = V_{11} \pm V_{12}, \quad (i = 1, 2), \tag{18}$$

the + and − signs corresponding to the symmetric ($i = 1$) and antisymmetric ($i = 2$) wave functions respectively.

We see that the interaction V splits the unperturbed level $E^{(0)}$, equation (12), into two levels. We have shown these levels in *Figure 11* for a repulsive interaction, $V(1, 2) > 0$, so that $V_{11} > 0$. For the purpose of illustration we have also assumed $V_{11} > V_{12} > 0$.

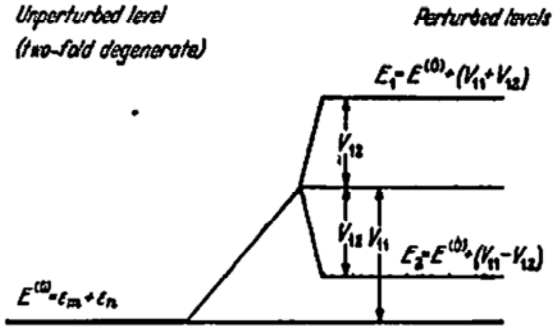
Unperturbed level
(two-fold degenerate)
Perturbed levels
$E_1 = E^{(0)} + (V_{11} + V_{12})$
V_{12}
V_{12}
V_{11}
$E_2 = E^{(0)} + (V_{11} - V_{12})$
$E^{(0)} = \varepsilon_m + \varepsilon_n$

Figure 11. Splitting and displacement of degenerate energy level of two identical particles by their interaction, in first-order perturbation theory

The case $m = n$ represents an exception. The energy level

$$E^{(0)} = 2\varepsilon_n \tag{19}$$

is not degenerate, the only eigenfunction being

$$\xi_1 = u_n(1)u_n(2). \tag{20}$$

The effect of the perturbation V is to shift the energy level (19) to

$$E = E^{(0)} + E^{(1)} = E^{(0)} + (\xi_1, V\xi_1) = E^{(0)} + V_{11}, \tag{21}$$

as shown in *Figure 12*, again for a repulsive interaction.

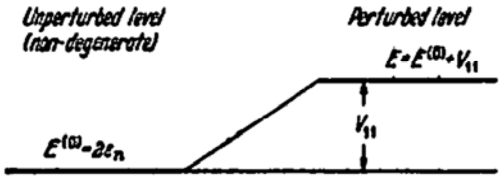
Unperturbed level
(non-degenerate)
Perturbed level
$E = E^{(0)} + V_{11}$
V_{11}
$E^{(0)} = 2\varepsilon_n$

Figure 12. Displacement of non-degenerate energy level of two identical particles by their interaction, in first-order perturbation theory

The matrix element V_{11} represents the mean interaction of the two particles in the states u_m and u_n. As such, we expect it to supplement the zero-order energy of the system. The occurrence of the additional term V_{12} in (18) on the other hand cannot be understood in this classical manner. It is a direct consequence of the indistinguishability of the two particles and represents a quantum-mechanical exchange interaction. Hence, terms of this type are often called *exchange integrals.* The occurrence of these exchange interaction terms is of very great importance, particularly for the understanding of the homopolar or covalent chemical bond.

We now take the two identical particles, which we are considering, to be electrons. In that case we know from section 28 that the correctly antisymmetrized zero-order wave functions, allowing for spin, which are derived from (16) are

$$\begin{aligned}\varphi_1^{(0)}\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1)\\ &=\frac{1}{\sqrt{2}}\{u_m(1)u_n(2)+u_m(2)u_n(1)\}\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1), \quad (22\text{a})\end{aligned}$$

and

$$\begin{aligned}\varphi_2^{(0)}\left\{\begin{matrix}\alpha_1\alpha_2\\ \frac{1}{\sqrt{2}}(\alpha_1\beta_2+\alpha_2\beta_1)\\ \beta_1\beta_2\end{matrix}\right\}\\ &=\frac{1}{\sqrt{2}}\{u_m(1)u_n(2)-u_m(2)u_n(1)\}\left\{\begin{matrix}\alpha_1\alpha_2\\ \frac{1}{\sqrt{2}}(\alpha_1\beta_2+\alpha_2\beta_1)\\ \beta_1\beta_2\end{matrix}\right\}. \quad (22\text{b})\end{aligned}$$

Since $H^{(0)}$ and V are spin-independent, introducing the spin wave functions does not affect the *position* or *splitting* of the perturbed energy levels, but it does affect the *multiplicity* of the levels. The levels E_1 [*Figure 11* and equation (17)] and E [*Figure 12* and equation (21)] arising from symmetric spatial wave functions are singlet levels, as given by equation (22a); the level E_2 [*Figure 11* and equation (17)] arises from an antisymmetric spatial wave function and hence leads to a three-fold degenerate triplet level, as given by equation (22b).

We consider the particular case of the first excited state of helium, which has a $(1s,\ 2p)$ configuration, *i.e.* one electron is in the ground state, $1s$, and one electron is in the first excited state, $2p$. On account of the Coulomb repulsion between the electrons, this 12-fold degenerate level splits into two levels; a 3-fold degenerate singlet level, specified by $S=0$, $L=1$, $\mathcal{J}=1$, and derived from the symmetric orbital functions (*cf* sections 25 and 27)

$$\xi_1^{(m)}=\frac{1}{\sqrt{2}}\{\Psi_1(0)\Psi_2(1,m)+\Psi_2(0)\Psi_1(1,m)\}, \quad (m=-1,0,1), \quad (23)$$

the total wave functions being

$$\xi_1^{(m)}\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1);$$

and a 9-fold degenerate triplet level specified by $S=1$ and $L=1$ (with all possible resultants $J=0$, 1 or 2) derived by combining the antisymmetric orbital functions

$$\xi_2^{(m)}=\frac{1}{\sqrt{2}}\{\Psi_1(0)\Psi_2(1, m)-\Psi_2(0)\Psi_1(1, m)\}, \quad (m=-1, 0, 1), \tag{24}$$

in all possible ways with the triplet spin functions $\alpha_1\alpha_2$, *etc.* That the energy levels with $m=-1, 0, 1$ are degenerate follows from the fact that we are considering a problem which is spherically symmetric whereas the value of m depends on the orientation of the coordinate axes. But for a spherically symmetric problem this orientation cannot be of physical significance. In Chapter IX we shall consider this rather general type of argument in greater detail.

Example 3: *Normal Zeeman effect*

In section 26 we mentioned the normal Zeeman effect, restricting ourselves to a single electron. We shall now consider an atom containing N electrons whose spin we shall neglect completely. Placing the atom in a homogeneous magnetic field H_z acting in the z-direction, gives rise to an interaction energy V, analogous to equation (26.1),

$$V=\frac{e}{2\mu c}H_zL_z, \tag{25}$$

where L_z is the z-component of the total orbital angular momentum $\mathbf{L}$ of the atom.

We know from section 25 that, if the Hamiltonian $H^{(0)}$ of the atom in the absence of the magnetic field H_z commutes with L_x, L_y and L_z, then the stationary states of the atom are eigenfunctions of $\mathbf{L}^2$ with the eigenvalues $\hbar^2n(n+1)$, $(n=0, 1, 2 \ldots)$*. The eigenvalue $\hbar^2n(n+1)$ is $(2n+1)$-fold degenerate, corresponding to the $(2n+1)$ eigenstates $\Psi(n, m)$, $(m=-n, -n+1, \ldots n)$ of L_z whose corresponding eigenvalues are $\hbar m$, $(m=-n, \ldots n)$.

We now want to calculate the effect of placing the atom in a magnetic field H_z on such a $(2n+1)$-fold degenerate energy level $E^{(0)}$. We do not have to use perturbation theory in this particular case. For the wave functions $\Psi(n, m)$ are exact eigenfunctions not only of the unperturbed Hamiltonian $H^{(0)}$ but also of the perturbation V:

$$(H^{(0)}+V)\Psi(n, m)=\left(E^{(0)}+\frac{e\hbar}{2\mu c}H_zm\right)\Psi(n, m), \quad (m=-n, \ldots n), \tag{26}$$

---

* In Chapter IX we shall see that the conditions

$$[H^{(0)}, L_x]=[H^{(0)}, L_y]=[H^{(0)}, L_z]=0$$

imply that the Hamilton $H^{(0)}$ of the atom, in the absence of the magnetic field, is spherically symmetric, *i.e.* it is invariant under rotations. This is certainly the case, if the potential energy term, entering $H^{(0)}$, depends only on the distances between the electrons, and between the electrons and the nucleus. (*Cf* the footnote referring to equation (25.9), where an analogous case is discussed.)

i.e. the magnetic field H_z splits the $(2n+1)$-fold degenerate level $E^{(0)}$ into $(2n+1)$ equidistantly spaced levels, the separation between two neighbouring levels being $(eh/2\mu c)H_z$. The quantity $(eh/2\mu c)$, which is the natural unit of magnetic moment to use in atomic problems, is called the Bohr magneton. The quantum number m is often known as the magnetic quantum number as it was first encountered in the above problem and is only of significance if the spherical symmetry of the system is disturbed by some external agency such as an applied magnetic field.

EXERCISES VII

VII.1. Carry out the analysis, analogous to that of section 29, which leads to equations (30.9) and (30.10).

VII.2. The Schrödinger equation of a one-dimensional oscillator is given in section 11, equation (11.3). Calculate the first-order perturbations of the energy levels if a perturbation $V(x)=ax^3$ is applied.

VII.3. Obtain the energy levels, according to first-order perturbation theory, of a one-dimensional oscillator, the perturbation being $V(x)=ax^4$.

VII.4. Obtain the eigenfunctions and eigenvalues of the perturbation problem defined in section 29 to the second approximation, *i.e.* including terms of the order of magnitude of ε^2, assuming the eigenvalues of the unperturbed Hamiltonian to be non-degenerate.

VII.5. If H is the diagonal matrix with eigenvalues E_1, E_2 and E_3 $(E_1 \neq E_2 \neq E_3 \neq E_1)$, use perturbation theory to find the eigenvalues and eigenvectors of the matrix $(H+\varepsilon V)$ to first order in the small quantity ε, V being a 3×3 matrix. Consider the same problem in the degenerate case $E_1=E_2\neq E_3$.

VIII

COLLISION PROCESSES

We now come to consider collision processes, *i.e.* processes in which an incident particle (such as an electron, proton or neutron) strikes a target (*e.g.* an atom or nucleus) which scatters it. The target may remain in its initial state (elastic collisions) or it may be 'knocked' into an excited state, emitting other particles if the excitation is sufficiently high (inelastic processes).

The theoretical analysis of such collision phenomena is of fundamental importance. For scattering experiments represent one of the best methods of studying the properties of atoms and nuclei, and the interactions of elementary particles with each other and with atomic and nuclear systems.

Only a very few collision problems are capable of exact analytic solutions in closed form. In general, these problems are very complicated and require special approximation methods suitable for the problem considered; even these often lead to very heavy computational analysis.

It is, of course, not our object to discuss these special methods of treating collision problems*. We shall deal only with some of the general principles underlying the theory of scattering and shall derive some of the simplest and most important results which often form the starting point for more elaborate approximation methods.

32. Elastic scattering by a fixed centre of force

We begin by considering the idealized problem of a beam of particles of mass μ travelling in the z-direction with velocity v, being scattered by a fixed scattering centre placed at the origin of coordinates, $\mathbf{r}=0$, and acting primarily in the neighbourhood of the origin. We describe this scattering centre by a potential $V(\mathbf{r})$. As a result of this potential field, the incident beam of particles experiences a force and is deflected. The scattering centre in an actual experiment will be an atom or a nucleus (or, rather, a whole group of atoms, *etc*, but for the purpose of analysis we need consider only one scattering centre). The potential which the incident particle experiences is then very much more complicated, depending not only on the separation $\mathbf{r}$ but also on the internal coordinates of the target. Nevertheless, one can in many cases represent the atomic or nuclear scattering centre approximately by such a potential. We have stated above that the scattering centre is fixed. This would correspond to an infinitely heavy atom or nucleus. But by interpreting μ as the reduced mass (*cf* Exercise III.3) and v as the velocity of the incident particle relative to the scattering centre, our considerations apply equally to a scattering centre of finite mass. We shall come back to this point in section 35.

---

* These questions are treated comprehensively in the book by Mott and Massey[22] and the articles by Wentzel[24] and Massey[25].

Mathematically, the problem we have to consider is the Schrödinger equation

$$\left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{1}$$

The potential $V(\mathbf{r})$, describing the scattering centre, will be most effective near the origin, but at large distances from the origin its effect will be small. More precisely, we shall assume that

$$rV(\mathbf{r}) \to 0 \quad (r = |\mathbf{r}|) \tag{2}$$

as $r \to \infty$, in any direction*. Under these conditions the total energy E of the particle is also equal to its kinetic energy at large distances from the scattering centre where $V(\mathbf{r}) = 0$, *i.e.*

$$E = \tfrac{1}{2}\mu v^2. \tag{3}$$

It is convenient to measure the potential in units of $\hbar^2/2\mu$, *i.e.* define

$$U(\mathbf{r}) = \frac{2\mu}{\hbar^2} V(\mathbf{r}), \tag{4}$$

and to put

$$k = \mu v/\hbar, \tag{5}$$

i.e. k is the magnitude of the wave vector of the incident particle. Equation (1) then becomes

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = U(\mathbf{r})\psi(\mathbf{r}). \tag{6}$$

On account of our assumptions about the potential, equation (2), equation (6) has a continuous spectrum of eigenvalues for $E > 0$, *i.e.* the kinetic energy of the incident particle (at large distances, *i.e.* outside the scattering field) may assume any positive value.

We now require a solution $\psi(\mathbf{r})$ of (6) which satisfies two conditions. Firstly, we want

$$\psi(\mathbf{r}) \to e^{ikz} \quad \text{as} \quad V(\mathbf{r}) \to 0 \tag{7}$$

i.e. if we decrease the strength of the scattering potential, the solution should tend to the plane wave solution representing a monochromatic beam of particles. Secondly, we want a solution representing the incident beam of particles, given by the plane wave

$$e^{ikz}$$

---

* This condition is not satisfied for the Coulomb potential $1/r$ which is a 'long-range' potential. Hence our considerations do not apply in this case. But in practice we do not usually deal with a pure Coulomb field, but with a screened Coulomb field, for which the potential decreases sufficiently rapidly. For example, the Coulomb field of the atomic nucleus is screened by the electrons surrounding it.

(which we have normed so that the incident beam of particles contains one particle per unit volume), together with outgoing waves which have been scattered from the target. From equation (12.12) we expect equation (6) to possess two asymptotic solutions

$$\frac{1}{r}e^{\pm ikr}f(\theta, \varphi), \tag{8}$$

representing outgoing and ingoing spherical waves of particles respectively, with the same energy $E=\hbar^2k^2/2\mu$ as the incident beam, confirming, as is physically obvious, that we are dealing with elastic scattering. Only the former solution is physically acceptable for our problem, so that the total asymptotic form of the solution of (6) which we require is

$$e^{ikz}+\frac{1}{r}e^{ikr}f(\theta, \varphi), \tag{9}$$

the first term representing the incident wave, the second the scattered wave. $f(\theta, \varphi)$ is called the scattering amplitude of the process.

To specify the angular distribution of the particles which have been scattered, we consider the passage of these particles through a large sphere of radius R whose centre is at the target. Consider an element of surface $dS=R^2d\Omega$, $d\Omega$ being an element of solid angle, on the sphere subtending the polar angles (θ, φ) at the target. The number of scattered particles crossing dS per unit time or, as we shall more briefly say, the current of particles scattered across dS, is given by the product (density of scattered particles at dS) × (velocity of scattered particles) × (element of area dS), *i.e.* by

$$\begin{pmatrix}\text{Current of particles}\\ \text{scattered across } dS\end{pmatrix}=\left|\frac{1}{R}e^{ikR}f(\theta, \varphi)\right|^2\left(\frac{\hbar k}{\mu}\right)R^2d\Omega$$

$$=\left(\frac{\hbar k}{\mu}\right)|f(\theta, \varphi)|^2d\Omega. \tag{10}$$

This expression gives the relative probabilities of a particle being scattered into two different directions. The absolute probability, however, still depends on the flux of the incident beam of particles striking the target. The current of particles scattered across dS is clearly proportional to the flux of incident particles, *i.e.* the number of particles incident per unit area per unit time. Since we have taken the density of the incident beam to be unity, the flux is simply equal to the velocity $(\hbar k/\mu)$. We then define the differential cross-section $\sigma(\theta, \varphi)d\Omega$ by the equation

$$\sigma(\theta, \varphi)d\Omega=\frac{(\text{Current of particles scattered across } dS)}{(\text{Flux of incident beam})}$$

i.e.

$$\sigma(\theta, \varphi)d\Omega=|f(\theta, \varphi)|^2d\Omega. \tag{11}$$

We see that the cross-section has the dimensions of an area. It is the effective area which the target presents to the incident beam of particles for the process considered.

From the angular distribution or differential cross-section (11) we obtain the total cross-section by integrating over all angles,

$$\sigma_{\text{tot}} = \int_{\substack{\text{unit}\\\text{sphere}}} |f(\theta, \varphi)|^2 d\Omega. \tag{12}$$

To find the scattering amplitude $f(\theta, \varphi)$, *i.e.* to solve equation (6), it is most convenient to convert it first to an integral equation. We shall not give a mathematically rigorous derivation of this integral equation, but only a plausible one which brings out its physical significance.

33. The integral equation of the scattering problem

Before considering the wave equation (32.6), let us consider an electrostatic problem with which the reader may be more familiar, namely, to obtain the potential $\psi(\mathbf{r})$ due to a distribution of electric charge given by the density function $\varrho(\mathbf{r})$. This potential is the solution of Poisson's equation

$$\nabla^2\psi(\mathbf{r}) = -4\pi\varrho(\mathbf{r}). \tag{1}$$

If we consider a unit charge at the point $\mathbf{r}=\mathbf{r}'$, this equation reduces to

$$\nabla^2\psi(\mathbf{r}) = -4\pi\delta(\mathbf{r}-\mathbf{r}'), \tag{2}$$

where $\delta(\mathbf{r}-\mathbf{r}')$ is the three-dimensional Dirac δ-function defined in equation (4.18) *. The solution of this equation is well known to be the Coulomb potential

$$\psi(\mathbf{r}) = \frac{1}{|\mathbf{r}-\mathbf{r}'|}, \tag{3}$$

as one can also easily verify from equation (2). To obtain the potential due to a general charge distribution we go back to equation (1).

* For the benefit of readers who omitted section 4 we briefly state the relevant properties of the Dirac δ-function $\delta(x)$. It is defined by

$$(i)\ \ \delta(x) = 0 \qquad \text{if } x \neq 0$$

$$(ii)\ \ \lim_{\epsilon\to 0} \delta(\epsilon) \to \infty \text{ in such a way that } \int \delta(x)dx = 1$$

for any range of integration enclosing the origin.

From this definition its most important property follows at once: if $a<b$, then for any function $f(x)$,

$$\int_a^b f(x)\delta(x-c)dx = \begin{cases} f(c), \text{ if } a<c<b \\ 0, \text{ if } c<a \text{ or } c>b. \end{cases}$$

The three-dimensional δ-function $\delta(\mathbf{r})$ is defined by

$$\delta(\mathbf{r}) = \delta(x)\delta(y)\delta(z)$$

and it follows at once that for any volume of integration V:

$$\int_V f(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}_0)d\mathbf{r} = \begin{cases} f(\mathbf{r}_0), \text{ if the point } \mathbf{r}_0 \text{ lies } \textbf{inside} \text{ the volume } V \\ 0 \quad ,\ \ ,, \quad ,, \quad \mathbf{r}_0\ ,,\ \textbf{outside} \quad ,, \quad ,, \end{cases}$$

Since this equation is linear we expect the potential it generates to be a linear superposition of the potentials due to the individual elements of charge $\varrho(\mathbf{r}')d\mathbf{r}'$ in the different volume elements $d\mathbf{r}'$, *i.e.*

$$\psi(\mathbf{r}) = \int \frac{\varrho(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|}, \tag{4}$$

the integral being over all space, and this expression represents the complete solution of our electrostatic problem.

We now treat the inhomogeneous wave equation

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = F(\mathbf{r}) \tag{5}$$

in the same way as equation (1). Whereas solutions of the homogeneous wave equation represent waves in the absence of the scattering target, we can think of $F(\mathbf{r})$ in (5) as a *source term* which generates or absorbs waves. The solution of equation (5) then consists of two parts. It consists of waves generated by the source term $F(\mathbf{r})$, to which may be added any solution of the homogeneous wave equation

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = 0,$$

representing waves generated and absorbed at infinity (*i.e.* at very large distances) but not affected by the source term $F(\mathbf{r})$.

We again start by considering a unit source at the origin, so that (5) becomes

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = \delta(\mathbf{r}). \tag{6}$$

We easily verify by direct substitution that equation (6) possesses the following two solutions, which are functions of the radial coordinate r only,

$$\psi(r) = -\frac{e^{\pm ikr}}{4\pi r}, \tag{7}$$

(*cf* Exercise VIII.1), corresponding to outgoing spherical waves generated by the unit source at the origin and to incoming spherical waves absorbed by the unit sink at the origin respectively. As we are interested in outgoing spherical waves only, in connection with scattering problems, we shall only require the solution containing the positive exponential function.

Analogously to the treatment of Poisson's equation, we can now at once write down a particular integral of the wave equation (5) containing a general source term $F(\mathbf{r})$, as a linear superposition of waves produced by different elements of the source. The solution corresponding to outgoing waves only is given by

$$\psi(\mathbf{r}) = -\frac{1}{4\pi}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}F(\mathbf{r}')d\mathbf{r}'. \tag{8}$$

To this particular integral of the wave equation we can still add any solution of the homogeneous wave equation.

We can apply these results to the scattering problem of section 32. From the discussion leading to equation (32.9), we see that the solution of the homogeneous wave equation which we must add to (8) is e^{ikz}. Furthermore, comparing equations (32.6) and (5) above, we see that $F(\mathbf{r}) = U(\mathbf{r})\psi(\mathbf{r})$, *i.e.* we finally obtain the following integral equation representing our scattering problem

$$\psi(\mathbf{r}) = e^{ikz} - \frac{1}{4\pi}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}', \tag{9}$$

the integration being over all space. The great advantage in expressing the scattering problem by this integral equation lies in the fact that this equation already incorporates all our boundary conditions.

We introduce the unit vector $\mathbf{n}$ in the direction of the scattered particle, defined by the polar angles θ, φ, *i.e.* $\mathbf{n}$ has the components $(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, and the wave vector $\mathbf{k}$ of the scattered particle, so that

$$\mathbf{k} = k\mathbf{n}. \tag{10}$$

To obtain the differential cross-section, we require the asymptotic form, for large values of r, of the wave function (9). Since the potential $U(\mathbf{r}')$ in (9) vanishes for large r', we can, for sufficiently large values of r, *i.e.* $r \gg r'$, put

$$|\mathbf{r}-\mathbf{r}'| \simeq r - \mathbf{n}\mathbf{r}', \quad \frac{1}{|\mathbf{r}-\mathbf{r}'|} \simeq \frac{1}{r}. \tag{11}$$

Substituting (11) into (9) leads to the asymptotic wave function for large r,

$$\psi(\mathbf{r}) \simeq e^{ikz} - \frac{e^{ikr}}{4\pi r}\int e^{-i\mathbf{k}\mathbf{r}'} U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}' \tag{12}$$

i.e. the wave function has the required asymptotic form as given by equation (32.9), with the scattering amplitude $f(\theta, \varphi)$ being defined by

$$f(\theta, \varphi) = -\frac{1}{4\pi}\int e^{-i\mathbf{k}\mathbf{r}'} U(\mathbf{r}')\psi(\mathbf{r}')d\mathbf{r}'. \tag{13}$$

Substituting this equation into (32.11) for the differential cross-section we obtain, using (32.4),

$$\begin{aligned} \sigma(\theta, \varphi)d\Omega &= \left(\frac{\mu}{2\pi\hbar^2}\right)^2 \left| \int e^{-i\mathbf{k}\mathbf{r}} V(\mathbf{r})\psi(\mathbf{r})d\mathbf{r} \right|^2 d\Omega \\ &= \left(\frac{\mu}{2\pi\hbar^2}\right)^2 |(e^{i\mathbf{k}\mathbf{r}}, V(\mathbf{r})\psi(\mathbf{r}))|^2 d\Omega. \end{aligned} \tag{14}$$

Equation (14) is a rigorous expression for the differential cross-section. The number of particles scattered into states with momentum $\hbar\mathbf{k}$ is seen to be *fully* determined by the matrix element of the scattering potential between the actual stationary state $\psi(\mathbf{r})$ of the system and the particular scattered state $e^{i\mathbf{k}\mathbf{r}}$ in which we are interested.

From equations (14) and (32.10–11), we obtain the current of particles scattered across dS, as

$$\begin{pmatrix}\text{Current of particles}\\ \text{scattered across } dS\end{pmatrix}=\frac{\hbar k}{\mu}\sigma(\theta,\varphi)d\Omega=\frac{\mu k d\Omega}{(2\pi)^2\hbar^3}|(e^{i\mathbf{k}'\cdot\mathbf{r}}, V(\mathbf{r})\psi(\mathbf{r}))|^2.$$

We can write this expression in the following form

$$\begin{pmatrix}\text{Current of particles}\\ \text{scattered across } dS\end{pmatrix}=\frac{2\pi}{\hbar}\int_{k'}\frac{k'^2dk'd\Omega}{(2\pi)^3}|(e^{i\mathbf{k}'\cdot\mathbf{r}}, V(\mathbf{r})\psi(\mathbf{r}))|^2\delta(E'-E) \quad (15)$$

the integration being over k' only, but not over the angular coordinates θ and φ. Here $E=\hbar^2k^2/2\mu$ is the energy of an incident particle with wave vector $\mathbf{k}=(0, 0, k)$, as before, and $E'=\hbar^2k'^2/2\mu$ is the energy of a scattered particle with wave vector $\mathbf{k}'=(k'\sin\theta\cos\varphi, k'\sin\theta\sin\varphi, k'\cos\theta)$. We seem to have ascribed a wave vector of magnitude k' to the scattered particle, different to that of the incident particle, but the conservation of energy has been ensured through the δ-function $\delta(E'-E)$ in equation (15), on account of which only those scattering processes are possible for which $E'=E$. Equation (15) is easily verified to be equivalent to (14), by carrying out the integration over k'.

We can now interpret (15) as follows. In a scattering experiment one uses counters to count the number of events of a particular kind; in our case, the number of particles scattered into a definite element of solid angle $d\Omega$ with a given energy E', *i.e.* the number of scattered particles with a wave vector in the interval $(\mathbf{k}', \mathbf{k}'+d\mathbf{k}')$. Though we are dealing with a stationary state, we can in a certain sense think of the scattering centre as causing transitions to the group of states with wave vectors in the interval $(\mathbf{k}', \mathbf{k}'+d\mathbf{k}')$ in which we are interested. Now

$$\frac{d\mathbf{k}'}{(2\pi)^3}=\frac{k'^2dk'd\Omega}{(2\pi)^3}$$

represents the number of states with momentum vectors in the interval $(\mathbf{k}', \mathbf{k}'+d\mathbf{k}')$, as follows from equation (10.13)*. Hence we can interpret

$$\frac{2\pi}{\hbar}|(e^{i\mathbf{k}'\cdot\mathbf{r}}, V(\mathbf{r})\psi(\mathbf{r}))|^2\frac{d\mathbf{k}'}{(2\pi)^3}\delta(E'-E) \qquad (16)$$

as the *transition probability per unit time* to the group of states with wave vector in the interval $(\mathbf{k}', \mathbf{k}'+d\mathbf{k}')$. The fact that we are dealing with a stationary state, finds its expression in the fact that the transition probability per unit time is independent of time. This result for transition probabilities per unit time, equation (16), can be generalized greatly, as we shall see in section 37 where we shall consider a different approach to collision problems.

---

* In section 10 we had normalized the wave functions in a volume L^3 and found $(L/2\pi)^3d\mathbf{k}$ as the number of states with momentum vector in $(\mathbf{k}, \mathbf{k}+d\mathbf{k})$. In this section we have taken $L=1$.

34. Born approximation

The expression which we derived for the differential cross-section in the last section, equation (33.14), does of course not yet represent a solution of the scattering problem, because it still contains the unknown wave function $\psi(\mathbf{r})$.

We will consider the special case where the scattering potential $V(\mathbf{r})$ is weak. It follows from equation (32.7) that the actual wave function $\psi(\mathbf{r})$ is in that case not much different from the incident plane wave e^{ikz}. We can think of the scattering potential $V(\mathbf{r})$ as a small perturbation which slightly distorts the plane wave.

Under these circumstances, we can, as a first approximation, replace the actual wave function $\psi(\mathbf{r})$ by the plane wave e^{ikz} in the matrix element in (33.14). This gives the very important first Born approximation, in which the differential cross-section (33.14) becomes replaced by

$$\sigma(\theta, \varphi)d\Omega = \left(\frac{\mu}{2\pi\hbar^2}\right)^2 \left|\int V(\mathbf{r})e^{i(\mathbf{k}_0-\mathbf{k})\mathbf{r}}d\mathbf{r}\right|^2 d\Omega$$

$$= \left(\frac{\mu}{2\pi\hbar^2}\right)^2 |(e^{i\mathbf{k}\mathbf{r}}, V(\mathbf{r})e^{i\mathbf{k}_0\mathbf{r}})|^2 d\Omega, \qquad (1)$$

where $\mathbf{k}_0 = (0, 0, k)$ is the wave vector of the incident particle so that

$$\mathbf{k}_0\mathbf{r} = kz.$$

If we look at the integral equation of the scattering problem, equation (33.9), we see that the first Born approximation corresponds to the first approximation of a solution by iteration, in which one replaces the true wave function $\psi(\mathbf{r})$ by the plane wave e^{ikz} under the integral sign. In other words, we assume that to a first approximation the wave scattered from the volume element $d\mathbf{r}'$ is the plane incident wave, unmodified by any scattered waves. In the Born approximation, the transition probability, equation (33.16), assumes a particularly clear meaning. The true wave function $\psi(\mathbf{r})$ is replaced by the 'unperturbed' plane wave e^{ikz} and we can think of the scattering potential $V(\mathbf{r})$ as causing transitions to other unperturbed states. There exist also higher Born approximations corresponding to iterative solutions of (33.9) of higher order, but these are usually too complicated to be of much practical use. The first Born approximation, however, is exceedingly simple and therefore very useful.

The momentum vector $(\mathbf{k}_0 - \mathbf{k})$ which occurs in the matrix element of the Born approximation, equation (1), represents the momentum (in units of $\hbar$) which the incident particle has transferred to the scattering target as a result of the collision. The matrix element itself is seen to be simply the Fourier transform (*cf* Chapter I, Exercise I.6) of the scattering potential.

The Born approximation (1) assumes a particularly simple form in the important case where the scattering potential is a function of the radial coordinate r only, $V = V(r)$. We take the momentum transfer vector

$$\mathbf{K} = \mathbf{k}_0 - \mathbf{k}$$

as polar axis of a new set of polar coordinates with polar angles α and β. Referred to these new coordinates, the vector $\mathbf{r}$ has rectangular components $(r \sin \alpha \cos \beta,\ r \sin \alpha \sin \beta,\ r \cos \alpha)$ and the vector $\mathbf{K}$ the components $(0, 0, K)$, so that

$$\mathbf{K}\mathbf{r} = Kr \cos \alpha.$$

The matrix element in (1) then becomes

$$\int_0^\infty V(r) r^2 dr \int_0^{2\pi} d\beta \int_0^\pi e^{iKr\cos\alpha} \sin \alpha \, d\alpha = 2\pi \int_0^\infty V(r) r^2 dr \int_{-1}^1 e^{iKrt} dt$$

$$= \frac{4\pi}{K} \int_0^\infty V(r) r \sin (Kr) dr. \qquad (2)$$

The scattering amplitude, equation (33.13), is now independent of φ and is only a function of θ, the angle through which the incident particle has been scattered. Using (2) the scattering amplitude becomes

$$f(\theta) = -\frac{2\mu}{\hbar^2 K} \int_0^\infty V(r) r \sin (Kr) dr \qquad (3)$$

and the cross-section is then obtained from equation (32.10). The only variable parameter entering the scattering amplitude $f(\theta)$, and therefore the cross-section, in the first Born approximation is the magnitude of the momentum transfer $\hbar K$. Since

$$\mathbf{K}^2 = (\mathbf{k}_0 - \mathbf{k})^2 = \mathbf{k}_0^2 + \mathbf{k}^2 - 2\mathbf{k}_0\mathbf{k} = 2k^2(1 - \cos \theta) = 4k^2 \sin^2 \frac{\theta}{2}, \qquad (4)$$

i.e.

$$K = 2k \sin \frac{\theta}{2}, \qquad (5)$$

$f(\theta)$ and the cross-section depend on the angle of scattering θ and the momentum of the incident particle $\hbar k$ through the combination $(k \sin \theta/2)$ only.

The validity of the Born approximation has been discussed by Mott and Massey[22]. They have shown that if one considers a scattering potential limited to a region of linear dimension a (say $r \leqslant a$) and whose strength is of the order of V_0, then the Born approximation is valid, provided that

$$\frac{V_0 a}{\hbar v} \ll 1, \qquad (6)$$

v being the velocity of the incident particle.

It follows from equation (6) that the Born approximation improves with increasing energy of the bombarding particle. For scattering of electrons by hydrogen or helium, the Born approximation is valid for bombarding energies above approximately 100 eV, for heavier atoms (with larger V_0) this minimum energy increases to about 1 keV. For the scattering of neutrons or protons by nuclei, Born approximation begins to give moderately adequate results at 100 to 150 MeV, but is not really satisfactory below 200 MeV.

Equation (6) is a sufficient condition for the validity of the Born approximation, but often it is too severe a restriction, *i.e.* the Born approximation gives good results even though (6) fails to hold. In particular, the Born approximation may give small angle scattering quite well, but break down for large angle scattering. For small angle scattering is due to particles moving past the scatterer at great distances, whereas for large angle deflections the projectile must penetrate deeply into the scattering centre.

We consider two simple examples of the Born approximation.

Example 1: *Scattering by a square well potential*

Take the scattering potential $V(\mathbf{r})$ to be

$$\begin{aligned} V(r) &= -V_0, \quad r<a \\ &= 0, \qquad r>a \end{aligned}$$

then equation (3) gives at once

$$f(\theta)=\frac{2\mu V_0}{\hbar^2 K}\int_0^a r\sin(Kr)dr=\frac{2\mu V_0}{\hbar^2 K^3}\{\sin(aK)-aK\cos(aK)\}, \tag{7}$$

where θ and K are related by equation (5),

$$K=2k\sin\frac{\theta}{2},$$

and the differential cross-section is given by (32.11),

$$\sigma(\theta,\varphi)=\sigma(\theta)=|f(\theta)|^2.$$

Example 2: *Scattering by a screened Coulomb potential*

As the second example we will take the potential

$$V(r)=-\frac{Ze^2}{r}e^{-\alpha r} \tag{8}$$

corresponding to the scattering of an electron by an atom of nuclear charge Ze. The potential (8) represents the Coulomb field of the nucleus, being screened by the atomic electrons surrounding it. The parameter α determining the screening is approximately given by

$$\alpha=\frac{me^2Z^{\frac{1}{3}}}{\hbar^2}.$$

We obtain for the scattering amplitude

$$f(\theta)=\frac{2\mu Ze^2}{\hbar^2 K}\int_0^\infty e^{-r\alpha}\sin(Kr)\,dr=\frac{2\mu Ze^2}{\hbar^2(K^2+\alpha^2)}, \tag{9}$$

for the differential cross-section

$$\sigma(\theta)=|f(\theta)|^2=\left(\frac{2\mu Ze^2}{\hbar^2}\right)^2\frac{1}{(K^2+\alpha^2)^2}, \tag{10}$$

and for the total cross-section, on account of (4),

$$\begin{aligned}\sigma_{\text{tot}}&=2\pi\int_0^\pi\sigma(\theta)\sin\theta\,d\theta\\&=2\pi\int_0^{2k}\sigma(\theta)\frac{KdK}{k^2}\\&=\frac{2\pi}{k^2}\left(\frac{2\mu Ze^2}{\hbar^2}\right)^2\int_0^{2k}\frac{KdK}{(K^2+\alpha^2)^2}\\&=\frac{16\pi\mu^2Z^2e^4}{\hbar^4\alpha^2(\alpha^2+4k^2)}.\end{aligned} \tag{11}$$

If we consider the case of large momentum transfers, $K\gg\alpha$, we can neglect α^2 compared with K^2 in equations (9) and (10). Equation (10) for the differential cross-section in this case becomes, with $\hbar k=\mu v$,

$$\sigma_c(\theta)=\left(\frac{Ze^2}{2\mu v^2}\right)^2\operatorname{cosec}^4\frac{\theta}{2}, \tag{12}$$

which is Rutherford's celebrated formula for the scattering by a pure Coulomb potential $-Ze/r$. That we obtain this limiting formula agrees with the fact that for large momentum transfers the electron must penetrate well inside the atom so that the screening is no longer effective.

The differential cross-section (12) for Coulomb scattering in the forward direction ($\theta=0$) becomes infinite. This is due to the fact that the Coulomb potential is a long range potential which decreases only slowly with distance. Hence electrons passing even at a great distance from the atom experience small angle scattering.

To apply the validity criterion (6) to the Coulomb scattering, we can take the range and strength of the potential to be of the order of

$$a\simeq 1/\alpha,$$

$$V\simeq-\frac{Ze^2}{a}e^{-\alpha a}\Big|_{a=1/\alpha}=-Ze^2\alpha/e,$$

so that the validity criterion (6) becomes

$$\frac{Ze^2}{\hbar v}\ll 1.$$

Since $(e^2/\hbar c)$ is dimensionless and has the value 1/137 very nearly we can write this criterion as

$$\frac{Z}{137}\left(\frac{c}{v}\right) \ll 1.$$

This criterion also applies in the relativistic case, when $v/c \simeq 1$ (particle velocity approaches the velocity of light). It then only fails for the heaviest atoms, and even for these $Z/137 < 1$. This application of the validity criterion is of course in no way rigorous and is only meant to make the results (which are correct) appear plausible.

35. Laboratory and centre of mass coordinate systems

For the purpose of describing an experiment, one chooses a coordinate system corresponding as closely as possible to the experimental set-up, *i.e.* one considers the target initially at rest. This coordinate system, which we shall specify more fully in a moment, is known as the *laboratory coordinate system*. A theoretically much simpler description is obtained if we choose a coordinate system in which the centre of mass of the whole system (target + incident particle) is at rest. For in this system, known as the *centre-of-mass coordinate system*, the total momentum is always zero (both before and after collision).

We must now show how these two modes of description are related to each other.

In the laboratory coordinate system, we consider a particle of mass m which is shot at a target with velocity v. The target particle of mass m' is initially at rest, $v'=0$. As a result of the collision, the incident particle is scattered into a direction with polar angles θ_l, φ_l with velocity w, the target experiencing a recoil, in the direction θ_l', φ_l' with velocity w'. *Figure 13* (a) illustrates the laboratory coordinate system. The laws of conservation of energy and momentum give us the following four equations

$$\tfrac{1}{2}mv^2 = \tfrac{1}{2}mw^2 + \tfrac{1}{2}m'w'^2$$
$$mv = mw \cos \theta_l + m'w' \cos \theta_l'$$
$$mw \sin \theta_l = m'w' \sin \theta_l'$$
$$\varphi_l' = \varphi_l + \pi$$

The last of these equations simply states that a single process occurs entirely in one plane (so that we can arbitrarily put $\varphi_l=0$, $\varphi_l'=\pi$). The other three equations enable us to solve the scattering problem completely, given one of the final coordinates. For example, given θ_l (in addition to m, m', v; and $v'=0$) we can determine θ_l', w and w'. To determine θ_l itself as well, we would require the law of force governing the scattering process. Hence we can define a cross-section in the laboratory system,

$$\sigma_l(\theta_l, \varphi_l)d\Omega_l = \sigma_l(\theta_l, \varphi_l) \sin \theta_l d\theta_l d\varphi_l,$$

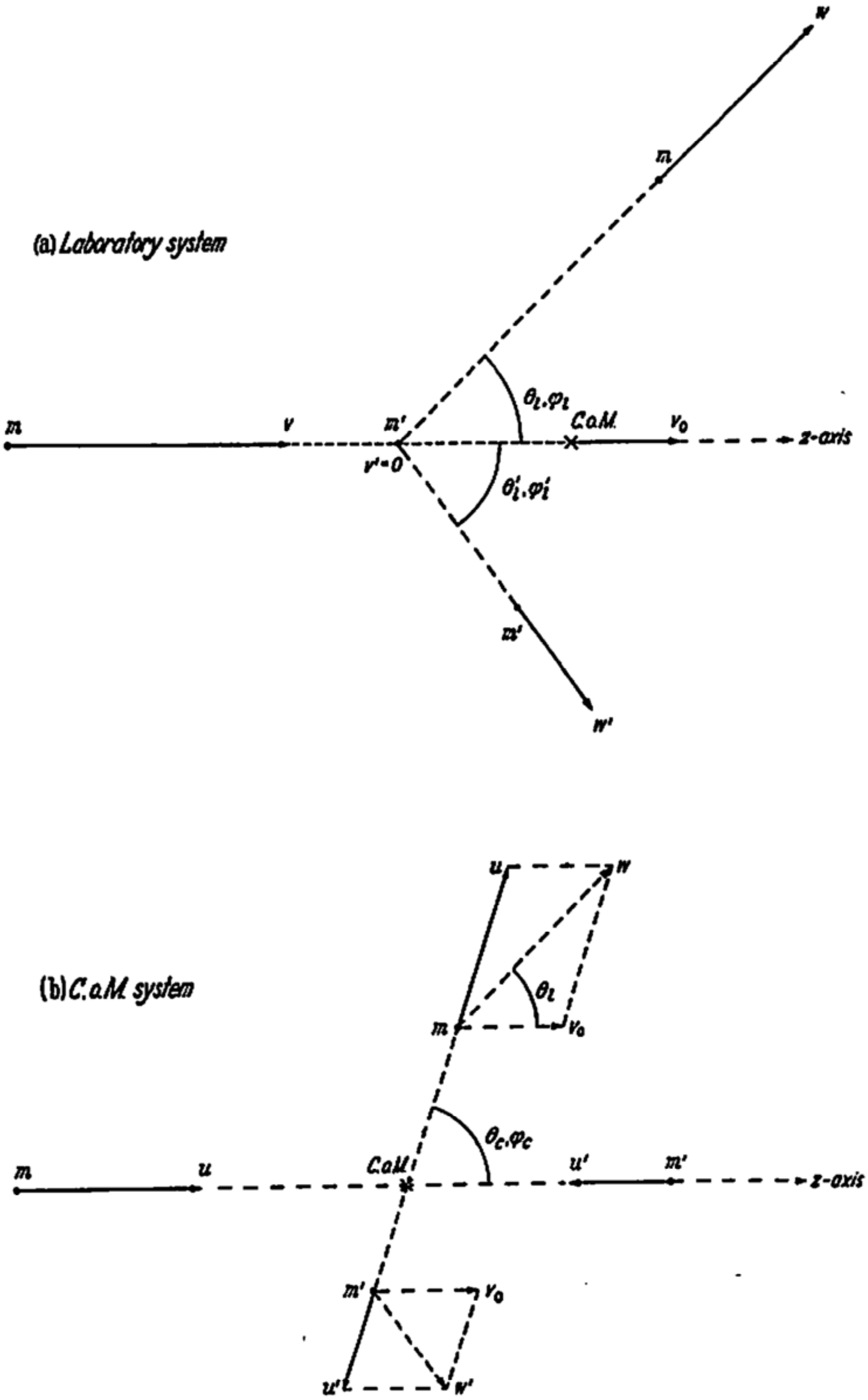
(a) Laboratory system
w
m
θ_l, φ_l
m
v
m'
C.a.M.
v_0
z-axis
$v'=0$
θ_l', φ_l'
m'
w'
(b) C.a.M system
u
w
θ_l
m
v_0
θ_c, φ_c
m
u
C.a.M
u'
m'
z-axis
m'
v_0
u'
w'

Figure 13. Description of scattering process in
(a) *laboratory coordinate system,*
(b) *centre-of-mass coordinate system.*

for scattering leading to a definite final state, analogously to the treatment in section 32*.

Just as in classical mechanics, so we can, in quantum mechanics too, impose a uniform velocity on our coordinate system (or the physical system), and these different coordinate systems are all equally good and must all give consistent descriptions of our system.

We will now go over to the centre-of-mass coordinate system in which the centre of mass (C.o.M.) of the system is permanently at rest. In the laboratory coordinate system, the centre of mass moves with a uniform velocity

$$v_0 = \frac{mv}{m+m'}.$$

We go over to the C.o.M. coordinate system by imposing a velocity, equal and opposite to v_0, on the system. We show this in *Figure 13* (b).

In the C.o.M. system, the scattering process can be thought of as two beams of particles travelling towards the C.o.M., itself at rest, being scattered through the same angle θ_c and moving off in opposite directions. The velocities of 'projectile' m and 'target' m' (which are now treated on an exactly equal footing) before collision are

$$u = v - v_0 = \frac{m'v}{m+m'}$$

$$u' = v' - v_0 = \frac{-mv}{m+m'},$$

(and $mu + m'u' = 0$, as required). The velocities after scattering through the angle θ_c are unchanged. The velocities in the two coordinate systems are related by these equations, which are evident from *Figure 13* (b):

$$\left.\begin{aligned} \varphi_l &= \varphi_c, \\ w \sin \theta_l &= u \sin \theta_c \\ w \cos \theta_l &= u \cos \theta_c + v_0. \end{aligned}\right\} \qquad (1)$$

From the last two equations, we obtain

$$\tan \theta_l = \frac{\sin \theta_c}{\cos \theta_c + a}, \qquad (2)$$

where

$$a = \frac{v_0}{u} = \frac{m}{m'}.$$

We can now define a differential cross-section in the C.o.M. system

$$\sigma_c(\theta_c, \varphi_c) d\Omega_c = \sigma_c(\theta_c, \varphi_c) \sin \theta_c d\theta_c d\varphi_c$$

for the colliding particles to be scattered through an angle θ_c, into the element of solid angle $d\Omega_c$. Now a particular element of solid angle

---

[*] In general the cross-section depends only on θ_l and is independent of φ_l. For it to depend on φ_l, the scattering potential must not be axially symmetric about the z-axis.

$d\Omega_l$ in the laboratory system will subtend a corresponding solid angle $d\Omega_c$ in the C.o.M. system, in such a way that

$$\sigma_l(\theta_l, \varphi_l)d\Omega_l = \sigma_c(\theta_c, \varphi_c)d\Omega_c, \tag{3}$$

for either side of this equation gives the current of particles through the same surface element per unit flux of incident beam.

We must now find the element of solid angle $d\Omega_l$ in the laboratory system corresponding to a particular element of solid angle $d\Omega_c$ in the C.o.M. system. From (1) we have that

$$d\varphi_l = d\varphi_c, \tag{4}$$

and from (2) we find that

$$\cos\theta_l = \frac{a + \cos\theta_c}{(1 + a^2 + 2a\cos\theta_c)^{\frac{1}{2}}},$$

so that

$$\sin\theta_l d\theta_l = \frac{(1 + a\cos\theta_c)}{(1 + a^2 + 2a\cos\theta_c)^{\frac{3}{2}}}\sin\theta_c d\theta_c \tag{5}$$

Equation (3) then gives for the relation of the cross-sections in laboratory and C.o.M. systems

$$\underline{\sigma_l(\theta_l, \varphi_l) = \sigma_c(\theta_c, \varphi_c)\frac{(1 + a^2 + 2a\cos\theta_c)^{\frac{3}{2}}}{|1 + a\cos\theta_c|}}. \tag{6}$$

Though the differential cross-sections differ in the two cases, the total cross-section derived by integrating over all directions in either coordinate system is of course the same. For the total cross-section is proportional to the total number of scattering events irrespective of direction of scattering.

We obtain a particularly simple important case by considering two particles of equal mass:

$$m' = m, \quad a = 1.$$

Equation (2) then gives

$$\theta_l = \theta_c/2;$$

as θ_l goes from 0 to $\pi/2$, θ_c goes from 0 to π. We obtain, with $\theta_c = 2\theta_l = 2\theta$ and $\varphi_c = \varphi_l = \varphi$,

$$\frac{\sigma_l(\theta, \varphi)}{\sigma_c(\theta, \varphi)} = \frac{\sin 2\theta d(2\theta)}{\sin\theta d\theta} = 4\cos\theta,$$

so that

$$\sigma_l(\theta, \varphi) = 4\cos\theta\ \sigma_c(\theta, \varphi),$$

as follows also from (6).

It follows from equation (2) for $a < 1$, *i.e.* the projectile is lighter than the target, that θ_l varies from 0 to π, as θ_c varies from 0 to π. For $a = 1$, *i.e.* for two particles of equal mass, we obtain the special case just considered. If $a > 1$, *i.e.* the projectile is heavier than the target, conservation of energy and momentum considerations restrict the possible scattering angles θ_l in the laboratory system. As θ_c varies from 0 to $\cos^{-1}(-1/a)$, θ_l increases from 0 to its maximum value

$\sin^{-1}(1/\alpha)$ which is always less than $\pi/2$. As θ_c increases further to π, θ_l decreases again to 0. We also see this from equation (5): if $\alpha > 1$, $(1 + \alpha \cos \theta_c)$ and hence $d\theta_l/d\theta_c$ are negative for $\theta_c > \cos^{-1}(-1/\alpha)$. In equation (6) we have introduced a modulus sign in the denominator as cross-sections are essentially positive quantities. The values of θ that occur for a given θ can be distinguished by the energy of the scattered particle,

$$\tfrac{1}{2}mw^2 = \tfrac{1}{2}m(v_0{}^2 + u^2 + 2v_0 u \cos \theta_c),$$

the smaller angle corresponding to the larger energy.

In the limiting case where one of the scattering particles is infinitely heavy, $m' \to \infty$ (in practice this means $m \ll m'$), this particle remains at rest all the time and coincides with the C.o.M. The two coordinate systems then become identical, as follows also at once from the above analysis. We considered this case in section 32 (where we wrote μ for the mass of the projectile which we now denote by m, and θ for the scattering angle). But we also saw in section 32 that we can interpret μ as the reduced mass of the incident particle with respect to the scattering centre, and we must now see how the cross-section derived in this way is related to the laboratory or C.o.M. cross-sections.

We saw in section 32 that, if one uses the reduced mass, one at the same time uses coordinates relative to the scattering centre, *i.e.* one uses a coordinate system (called the relative coordinate system) 'fixed' to the target particle, so that one calculates as though the target were infinitely heavy and consequently remained at rest. One easily sees that in this relative coordinate system the scattering angles θ_r and φ_r are identical with the corresponding angles in the C.o.M. system,

$$\theta_r = \theta_c, \quad \varphi_r = \varphi_c. \tag{7}$$

In order to see this, we reduce the target m' to rest in the C.o.M. system [*Figure 13* (b)] by imposing the appropriate velocity (namely, minus that of m' in the C.o.M. system) on the system, after the collision. This does not alter the direction but only the magnitude of the velocity of the scattered particle, so that (7) follows. Hence the cross-section calculated in the relative coordinate system is the C.o.M. cross-section.

In using these three coordinate systems one must use the appropriate energies. The energy in the laboratory system is

$$E_l = \tfrac{1}{2}mv^2,$$

the energy associated with the motion of the centre of gravity being

$$E_0 = \tfrac{1}{2}(m + m')v_0{}^2 = \frac{m}{m + m'}E_l.$$

The kinetic energy of the two particles in the C.o.M. system is

$$E_c = \tfrac{1}{2}mu^2 + \tfrac{1}{2}m'u'^2 = \frac{m'}{m + m'}E_l,$$

so that

$$E_t = E_0 + E_n$$

as required. For the energy of the incident particle, in the relative coordinate system, we have

$$E_r = \tfrac{1}{2}\mu v^2 = \tfrac{1}{2}\frac{mm'}{m+m'}v^2 = \frac{m'}{m+m'}E_l = E_c.$$

36. Inelastic scattering processes

We now come to consider inelastic scattering processes in which the state of the target is changed by the collision. The same stationary theory which was developed in sections 32 to 34 can easily be adapted to this case, so that we need only consider it briefly.

At first sight it may seem surprising that a time-independent approach can be used to describe these inelastic processes in which a collision changes the state of the target. What we really consider is, of course, the statistics of a succession of individual collisions between a single incident particle and a single atomic or nuclear scattering centre. We can best visualize this in the C.o.M. coordinate system, where we are dealing with two streams of particles (the 'incident' particles and the 'target' systems) colliding and flying off in opposite directions again, the scattered target being in a new state. (This treatment does of course not allow for such processes as double collisions of incident particles with the target. Correspondingly one must use a thin target, or modify the theory.)

Suppose the target system (*i.e.* the atom or nucleus) by itself is described by the Hamiltonian $H(q)$, q standing for a whole set of coordinates describing the target, and assume the complete sets of eigenfunctions and eigenvalues

$$\varphi_0(q), \varphi_1(q), \ldots \varphi_n(q), \ldots$$
$$E_0, E_1, \ldots E_n \ldots$$

of $H(q)$ are known. (For simplicity, we have here assumed a discrete spectrum of eigenvalues, but the generalization to the continuous or partly continuous case is easy.) If μ is the reduced mass of the incident particle P with position vector $\mathbf{r}=(x, y, z)$ relative to the C.o.M. of the scatterer, and $V(\mathbf{r}, q)$ is the interaction potential of P with the scatterer, the total system is described by the wave equation

$$\left\{\frac{-\hbar^2}{2\mu}\nabla^2 + H(q) + V(\mathbf{r}, q)\right\}\psi(\mathbf{r}, q) = E\psi(\mathbf{r}, q). \qquad (1)$$

We assume that the scattering potential is a short range potential, so that, analogously to equation (32.2),

$$rV(\mathbf{r}, q) \to 0 \quad \text{as} \quad r \to \infty, \qquad (2)$$

in any direction. If $\mathbf{k}_0 = (0, 0, k_0)$ is the wave vector of the incident particle, and $\varphi_0(q)$ the state of the scatterer before collision, then

$$E = \frac{1}{2\mu}\hbar^2 k_0^2 + E_0. \tag{3}$$

We again impose two conditions on the solutions of (1) we seek. Firstly, we require, analogously to equation (32.7), that

$$\psi(\mathbf{r}, q) \rightarrow e^{i\mathbf{k}_0\mathbf{r}}\varphi_0(q) \quad \text{as} \quad V(\mathbf{r}, q) \rightarrow 0. \tag{4}$$

For in the limit of vanishing interaction, we want the solution to describe a definite state of the target system, together with a plane wave, describing the uniform motion of the incident particles. Secondly, we require that the only incident waves are those corresponding to the beam of incident particles, all other waves being outgoing waves corresponding to scattered waves. This leads to a restriction on the asymptotic form of the solution analogous to (32.9) but more general, allowing for inelastic collisions. We can expand $\psi(\mathbf{r}, q)$ in terms of the complete set $\varphi_m(q)$, the coefficients being functions of $\mathbf{r}$ of course,

$$\psi(\mathbf{r}, q) = \sum_m \varphi_m(q)\psi_m(\mathbf{r}). \tag{5}$$

Substitution in (1) gives

$$\sum_m \varphi_m(q)\left\{\nabla^2 + \frac{2\mu}{\hbar^2}(E - E_m)\right\}\psi_m(\mathbf{r}) = U(\mathbf{r}, q)\psi(\mathbf{r}, q) \tag{6}$$

with

$$U(\mathbf{r}, q) = \frac{2\mu}{\hbar^2}V(\mathbf{r}, q), \tag{7}$$

and taking the inner product with $\varphi_s(q)$, just as far as the coordinates q of the scatterer are concerned, gives

$$(\nabla^2 + k_s^2)\psi_s(\mathbf{r}) = \int dq\overline{\varphi_s(q)}\,U(\mathbf{r}, q)\psi(\mathbf{r}, q) \equiv F_s(\mathbf{r}), \ (s = 0, 1, 2, \ldots), \tag{8}$$

where

$$k_s^2 = \frac{2\mu}{\hbar^2}(E - E_s), \quad (s = 0, 1, 2, \ldots). \tag{9}$$

Equations (8) and (9) describe collisions in which the target system after collision is in the state $\varphi_s(q)$. For $s = 0$, these are the elastic processes already considered in sections 32 to 34; for $s \neq 0$, they represent inelastic processes in which the scattered particle is described by a wave vector $\mathbf{k}_s = (k_s \sin\theta\cos\varphi, k_s\sin\theta\sin\varphi, k_s\cos\theta)$, whose magnitude is defined by the conservation of energy equation (9). Since the kinetic energy of the scattered particle must be positive, collisions leading to the state $\varphi_s(q)$ are only possible if

$$E_s < E = E_0 + \frac{\hbar^2}{2\mu}k_0^2.$$

The methods of section 33 at once give the set of integral equations equivalent to the set of equations (8):

$$\psi_s(\mathbf{r}) = e^{ik_0z}\delta_{0s} - \frac{1}{4\pi}\int\frac{e^{ik_s|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}F_s(\mathbf{r}')d\mathbf{r}', \quad (s=0, 1, \ldots). \tag{10}$$

Here the second term represents the scattered outgoing waves (there are no terms with negative exponentials, since there are no incoming spherical waves). The first term is only different from zero for $s=0$ and corresponds to the plane incident wave which occurs in the solution of the elastic scattering problem [*cf* equation (33.9)]. The asymptotic solution of (10) is

$$\psi_s(\mathbf{r}) \simeq e^{ik_0z}\delta_{0s} - \frac{e^{ik_sr}}{4\pi r}\int e^{-ik_s\mathbf{r}'}F_s(\mathbf{r}')d\mathbf{r}', \quad (s=0, 1, \ldots) \tag{11}$$

giving the scattering amplitude

$$f_s(\theta, \varphi) = -\frac{1}{4\pi}\int e^{-ik_s\mathbf{r}'}F_s(\mathbf{r}')d\mathbf{r}' \quad (s=0, 1, \ldots) \tag{12}$$

corresponding to the asymptotic wave form [*cf* equation (32.9)]

$$e^{ik_0z}\delta_{0s} + \frac{e^{ik_sr}}{r}f_s(\theta, \varphi), \quad (s=0, 1, \ldots). \tag{13}$$

Equations (12) and (13) refer to processes in which the target is left in the state $\varphi_s(q)$ after the collision. Analogously to section 32, we find for the differential cross-section for these processes

$$\sigma_s(\theta, \varphi)d\Omega = \frac{\left|\frac{1}{R}e^{ik_sR}f_s(\theta, \varphi)\right|^2\frac{\hbar k_s}{\mu}R^2d\Omega}{\left(\frac{\hbar k_0}{\mu}\right)} = \left(\frac{k_s}{k_0}\right)|f_s(\theta, \varphi)|^2d\Omega,$$
$$(s=0, 1, \ldots). \tag{14}$$

Apart from the factor (k_s/k_0) this agrees with equation (32.11).

By procedures entirely analogous to those of sections 33 and 34 we could consider transition probabilities and Born approximation. However, the reader should by this stage be well able to carry these investigations out for himself. The validity criterion for Born approximation, equation (34.6), must now be replaced by two analogous criteria for both incident and scattered velocities of the incident particle

$$\frac{V_0a}{\hbar v_0} \ll 1, \qquad \frac{V_0a}{\hbar v_s} \ll 1. \tag{15}$$

Amongst the many important problems in collision theory, which we have not discussed, that of collisions involving identical particles is particularly important. For the elastic scattering between two identical particles, one can easily show that the differential cross-section in the C.o.M. system is given by

$$\sigma(\theta, \varphi) = |f(\theta, \varphi) \pm f(\pi - \theta, \varphi + \pi)|^2 \tag{16}$$

where the ± signs correspond to states which are respectively symmetric and antisymmetric in the spatial coordinates.

In the case of inelastic scattering, it is less easy to allow for the identity of particles. In general, there exists some ambiguity as to how the wave functions should be symmetrized or antisymmetrized*. But in many collision problems involving fast particles, our above treatment is adequate; *e.g.* we can think of a fast electron, scattered by an atom so as to leave the atom with a high energy, as having retained its identity—it did not spend sufficiently long in the atom for electron exchange to have taken place.

37. Time-dependent approach to collision problems

So far we have treated collision problems by time-independent methods, *i.e.* we have looked for stationary states of a Hamiltonian $H = H^0 + V$, V describing a scattering potential which causes a distortion of the plane waves.

In practice we do, of course, not deal with infinitely extended plane waves. Rather, at some time $t = 0$ we switch on a particle accelerator, or open a shutter, allowing a beam of particles to strike a target. But we can make the wave packet sufficiently long for end effects to be unimportant so that the actual state can be approximated well by a stationary state.

Another approach to collision problems is by means of time-dependent perturbation theory. We shall, to begin with, develop this quite generally, and then come back to collision problems.

Given the Hamiltonian of a system to be

$$H(q, t) = H^0(q) + V(q, t), \tag{1}$$

the symbol q denoting the set of spatial (and spin) coordinates, we want to solve the time-dependent equation

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{2}$$

for the wave function $\psi(q, t)$, given its value at some instant $t = 0$: $\varphi(q, 0)$.

Suppose the unperturbed Hamiltonian H^0 possesses the complete orthonormal sets of eigenfunctions $\varphi_0(q), \varphi_1(q), \ldots \varphi_\nu(q), \ldots$ and eigenvalues $E_0, E_1, \ldots E_\nu, \ldots$, *i.e.*

$$H^0\varphi_\nu = E_\nu\varphi_\nu. \tag{3}$$

---

* For a full discussion, see Massey and Burhop, *Electronic and Ionic Impact Phenomena*, Oxford, Clarendon Press, 1952, p. 142 *et seq.*

We can then expand $\psi(q, 0)$ in terms of this set

$$\psi(q, 0) = \sum_{\nu=0}^{\infty} c_\nu(0)\varphi_\nu(q) \tag{4}$$

where we have denoted the constant expansion coefficients by $c_\nu(0)$, for reasons which will shortly be apparent*. We can expand the solution of the time-dependent problem (2) in terms of the time-dependent eigenfunctions $\varphi_\nu e^{-iE_\nu t/\hbar}$ of H^0:

$$\psi(q, t) = \sum c_\nu(t)\varphi_\nu(q)e^{-iE_\nu t/\hbar} \tag{5}$$

and we must now determine the time-dependent expansion coefficients $c_\nu(t)$. The physical interpretation of these coefficients is that $|c_\nu(t)|^2$ gives the probability of finding the system in the state $\varphi_\nu(q)e^{-iE_\nu t/\hbar}$ at time t.

Substituting (5) into (2) we obtain the following infinite system of equations for the coefficients $c_\nu(t)$:

$$i\hbar\frac{dc_\nu}{dt} = \sum_\mu V_{\nu\mu}c_\mu(t)e^{i(E_\nu - E_\mu)t/\hbar} \quad (\nu = 0, 1, \ldots) \tag{6}$$

where

$$V_{\nu\mu} = V_{\nu\mu}(t) = (\varphi_\nu(q), V(q, t)\varphi_\mu(q)) = \int dq\,\overline{\varphi_\nu(q)}\,V(q, t)\varphi_\mu(q). \tag{7}$$

We can solve the system of equations (6) by a method of successive approximations. We shall now assume that the perturbing potential V does not depend on the time explicitly. (The general case can also be easily derived.) We obtain the first approximation to the solution of (6) by putting

$$c_\nu(t) = c_\nu(0) \quad (\nu = 0, 1, \ldots)$$

on the right-hand side of equation (6). This gives at once

constant of integ

$$c_\nu(t) = c_\nu(0) - \sum_\mu V_{\nu\mu}c_\mu(0)\frac{e^{i(E_\nu - E_\mu)t/\hbar} - 1}{(E_\nu - E_\mu)}. \tag{8}$$

We shall, in particular, assume that initially, at time $t = 0$, the system is in an eigenstate of the unperturbed Hamiltonian H^0, which we take to be $\varphi_0(q)$, so that

$$c_\nu(0) = \delta_{\nu 0}. \tag{9}$$

(The general case, given by (4), can be deduced from this special case, which is, however, adequate for most applications.) Equations (8) then become

$$c_\nu(t) = \delta_{\nu 0} - V_{\nu 0}\frac{e^{i(E_\nu - E_0)t/\hbar} - 1}{(E_\nu - E_0)}, \quad (\nu = 0, 1, \ldots). \tag{10}$$

---

[*] We are writing eigenfunction expansions as sums over discrete states; if the problem has a continuous spectrum these summations are of course to be replaced by the appropriate integrals.

Substituting equations (10) in the right-hand side of (6) we can obtain a second approximation. But for the present we shall consider the first approximation only. This will be adequate if the perturbing potential V is weak or at any rate has acted for not too long a time interval t.

Since, with our assumptions, the system is initially in a definite eigenstate φ_0 of the unperturbed Hamiltonian H^0, we can now interpret $|c_\nu(t)|^2$ as the probability for a transition to the state φ_ν to have occurred in the time t, and we can think of V as causing these transitions.

In many problems, particularly in collision theory, we are dealing with continuous sets of states. We are then not interested in one individual state, but in groups of neighbouring states. The parameter ν now varies continuously. We shall accordingly write $E=E(\nu)$ instead of E_ν. We shall then consider the group of states for which the energy E lies in the interval

$$E-\Delta E<E(\nu)<E+\Delta E. \tag{11}$$

The probability for a transition to one of these states to have occurred in time t is then

$$\int d\nu|c_\nu(t)|^2=\int_{E-\Delta E}^{E+\Delta E} dE\frac{d\nu}{dE}|c_\nu(t)|^2=\frac{2\pi t}{\hbar}\int_{E-\Delta E}^{E+\Delta E} dE\frac{d\nu}{dE}|V_{\nu0}|^2 D\left(E-E_0,\frac{t}{2\hbar}\right), \tag{12}$$

where

$$D\left(E-E_0,\frac{t}{2\hbar}\right)=\frac{\sin^2[(E-E_0)t/2\hbar]}{\pi(E-E_0)^2 t/2\hbar}. \tag{13}$$

We considered a similar function in **Exercise 1.6.** For $t\gg\hbar/E_0$, it differs only appreciably from zero in the narrow interval

$$|E-E_0|<\hbar/t \tag{14}$$

having a sharp peak of height $t/\hbar$ at $E=E_0$. Hence, the transition probability to a state φ_ν is only appreciable if $E(\nu)\simeq E_0$, *i.e.* if energy is conserved, *within the limitations of the uncertainty principle.* Furthermore

$$\begin{aligned}\int_{-\infty}^{\infty} D\left(E-E_0,\frac{t}{2\hbar}\right)dE&=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2\left[(E-E_0)\frac{t}{2\hbar}\right]}{\left[(E-E_0)^2\frac{t}{2\hbar}\right]}dE\\&=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin^2 x}{x^2}dx\\&=1,\end{aligned}$$

independently of the value of $t/2\hbar$. Hence, if $(d\nu/dE)|V_{\nu0}|^2$ is a slowly varying function of E in the very narrow interval (14), compared with the variation of D, then the function D in the integrand of (12) acts simply as the δ-function $\delta(E-E_0)$, and we can write (12) in the form

$$\int d\nu |c_\nu(t)|^2 = \frac{2\pi t}{\hbar}\int dE \frac{d\nu}{dE}|V_{\nu 0}|^2\delta(E-E_0) \tag{15}$$

which is correct for any interval of integration. The transition probability (15) is proportional to the time. Hence we can define a transition probability per unit time, w, which is constant,

$$w = \frac{2\pi}{\hbar}\int dE \frac{d\nu}{dE}|V_{\nu 0}|^2\delta(E-E_0), \tag{16}$$

or, carrying out the integration,

$$w = \frac{2\pi}{\hbar}\left[\frac{d\nu}{dE}|V_{\nu 0}|^2\right]_{E=E_0}. \tag{17}$$

$(d\nu/dE)$ gives the number of states with energy in the range $(E, E+dE)$ per unit energy range. The form which our result takes in the most general case is now apparent. Suppose we label the eigenfunctions and eigenvalues of H^0 by the parameters $n_1, n_2, \ldots n_s$ which only assume discrete values (say 0, 1, 2, . . .), and by the continuous parameters $\nu_1, \nu_2, \ldots \nu_t$, *i.e.* we write these eigenfunctions and eigenvalues as $\varphi(n_1, \ldots n_s; \nu_1, \ldots \nu_t)$ and $E = E(n_1, \ldots n_s; \nu_1, \ldots \nu_t)$. Consider a particular initial state $\varphi_0 \equiv \varphi(a_1, \ldots a_s; \alpha_1, \ldots \alpha_t)$ with energy $E_0 \equiv E(a_1, \ldots a_s; \alpha_1, \ldots \alpha_t)$ and a group of final states for which the discrete parameters have the values $n_1, n_2, \ldots n_s$, and the continuous parameters lie in the intervals $(\nu_i, \nu_i + d\nu_i)$, $(i = 1, 2, \ldots t)$. Let $\varrho(n_1, \ldots n_s; \nu_1, \ldots \nu_t)d\nu_1 d\nu_2 \ldots d\nu_t$ be the number of states belonging to this group. Then the probability per unit time for a transition from the given initial state φ_0 to the group of final states considered, is

$$w = \frac{2\pi}{\hbar}\varrho(n_1, \ldots n_s; \nu_1, \ldots \nu_t)d\nu_1 d\nu_2 \ldots d\nu_t|V_{\nu 0}|^2\delta(E-E_0). \tag{18}$$

Equation (18) represents the most general form of our result. From it, any other case can be deduced by summing over appropriate ranges of the discrete parameters and integration over appropriate ranges of $\nu_1, \ldots \nu_t$. In particular, the transition probabilities obtained in section 33 and in Exercise VIII.2 are special cases of equation (18).

We must now discuss the physical meaning of the approximations we have made and see how far they are fulfilled in collision processes.

We assumed that, initially, at time $t=0$, the system was in an eigenstate φ_0 of the unperturbed Hamiltonian H^0, *i.e.* we treated the system as though the perturbation was not acting before $t=0$, and was only switched on at this instant. We have seen that an actual scattering experiment does not correspond to a truly stationary state; nor, however, does it correspond accurately to the type of time-dependent process we have considered. However, we made the assumption that $t \gg \hbar/E_0$. This means that we are considering intervals of time sufficiently long for 'end effects' (essentially, transients due to the sudden switching on of V at $t=0$ when the system is not in an eigenstate of H^0+V) to be unimportant. In practice, the condition $t \gg \hbar/E_0$ does not introduce any limitations, on account of the smallness of $\hbar$.

Secondly, we used first-order perturbation theory. This leads to transition probabilities proportional to the time, *i.e.* to stationary solutions (apart from the initial transients). It is also responsible for the matrix elements

$$V_{\nu 0} = (\varphi_\nu, V\varphi_0),$$

between the zero-order wave functions, occurring in equation (18). This is essentially equivalent to the first Born approximation (*i.e.* replacing the exact wave function $\psi(\mathbf{r})$ in (33.16) by the zero-order wave function $e^{i\mathbf{k}\cdot\mathbf{r}}$) for the general case under consideration. Hence we expect (18) to be valid under the same conditions, *i.e.* the potential V should be weak or the time t for which it acts should be sufficiently short. In collision problems, it again implies sufficiently high bombarding energies.

We shall not consider the higher approximations, but only state the result for the case where all the matrix elements $V_{\nu 0}$ in (18) vanish. In this case, transitions from the initial state to the group of final states of interest may still be possible in the second approximation. The transition probability is then still given by equation (18), with the matrix element $V_{\nu 0}$ replaced by

$$\sum_\mu \frac{V_{0\mu} V_{\mu\nu}}{(E_0 - E_\mu)}, \tag{19}$$

the summation being over the complete set of eigenfunctions of H^0. We see that transitions can occur in second-order approximation, if there exist states φ_μ for which $V_{0\mu}$ and $V_{\mu\nu}$ are both different from zero. Such states are called intermediate states.

In treating collision problems by stationary methods, we repeatedly stressed the asymptotic form which the wave function of the system must have, corresponding to a plane incident and spherical outgoing waves. By an analysis of the asymptotic form of the wave function (5), it can be shown that this wave function too has the correct asymptotic form (*cf* Wentzel[24], p. 736).

The reader should have no difficulties in applying the results of this section to particular scattering problems (*cf* Exercise VIII.5).

38. Partial waves

We now come to discuss briefly the general principles of a completely different approach to scattering problems which is of great importance in applications*. We discuss the same problem as in section 32: the scattering of a particle with mass, neglecting spin (*i.e.* either the particle is spinless or the interaction causing the scattering is spin-independent), by a spherically symmetric potential $V(r)$ which has a finite range a, *i.e.* for $r>a$,

$$V(r)=0. \tag{1}$$

This condition can be relaxed to include potentials which tend to zero faster than $1/r$ as $r\rightarrow\infty$.

In section 32 we considered a plane wave scattered by this potential. Now we shall analyse this plane wave into spherical waves and examine the distortion, due to the potential, of each separate partial spherical wave.

For this purpose we write the scattering equation (32.6) in spherical polar coordinates. For a spherically symmetric potential and spinless particles the problem is axially symmetric about the direction of motion of the incident particles which is taken as z-axis. Hence the solutions are independent of the azimuthal angle φ. Separating variables, as in sections 12 and 13, one finds solutions to equation (32.6) given by

$$G_l(r)P_l(\cos\theta), \quad (l=0, 1, 2\ldots) \tag{2}$$

where $G_l(r)$ is a solution of equation (12.7):

$$\left.\begin{array}{l}\dfrac{1}{r^2}\dfrac{d}{dr}\left(\dfrac{r^2dG_l(r)}{dr}\right)+\left[k^2-U(r)-\dfrac{l(l+1)}{r^2}\right]G_l(r)=0,\\ \qquad\qquad\qquad\qquad\qquad\qquad (l=0, 1, \ldots),\end{array}\right\} \tag{3}$$

for the interval $0<r<\infty$. Outside the range of the potential, *i.e.* for $r>a$, this equation reduces to the free particle equation which we shall consider first of all:

$$\left.\begin{array}{l}\dfrac{1}{x^2}\dfrac{d}{dx}\left(\dfrac{x^2dF_l(x)}{dx}\right)+\left[1-\dfrac{l(l+1)}{x^2}\right]F_l(x)=0,\\ \qquad\qquad\qquad\qquad\qquad (l=0, 1, \ldots),\end{array}\right\} \tag{4}$$

with

$$x=kr.$$

This is a second-order differential equation for each value of l and so possesses two linearly independent solutions $h_l^{(1)}(x)$ and $h_l^{(2)}(x)$, known as spherical Hankel functions of order l and of the first and second kind.

The spherical Hankel functions have been extensively investigated. We shall only state those properties which we require†.

---

* For comprehensive treatments *cf* Mott and Massey[22] and Massey[25]. There are also useful sections in Mott and Sneddon[10] and Schiff[12].

† For a collection of properties *cf* Schiff[12]; an excellent mathematical treatment of these functions is given by Sommerfeld[4]. The standard treatise is G. N. Watson's *Theory of Bessel functions*, Cambridge University Press, 1944, and Macmillan, New York, 1944.

The spherical Hankel functions of order l are related to the Hankel functions of order $l+\frac{1}{2}$ by

$$h_l^{(s)}(kr)=\left(\frac{\pi}{2kr}\right)^{\frac{1}{2}}H_{l+\frac{1}{2}}^{(s)}(kr), \quad (s=1, 2). \tag{5}$$

For real values of l and x (and we of course restrict ourselves to these), $h_l^{(1)}(x)$ and $h_l^{(2)}(x)$ are complex conjugates of each other:

$$\overline{[h_l^{(1)}(x)]}=h_l^{(2)}(x). \tag{6}$$

This follows in fact since equation (4) is a differential equation with real coefficients, so that if $h_l^{(1)}(x)$ is a solution so is its complex conjugate $h_l^{(2)}(x)$.

For large x, the $h_l^{(s)}(x)$ behave asymptotically like

$$h_l^{(1)}(x)\simeq\frac{1}{x}\exp\left\{i\left[x-(l+1)\frac{\pi}{2}\right]\right\} \tag{7a}$$

$$h_l^{(2)}(x)\simeq\frac{1}{x}\exp\left\{-i\left[x-(l+1)\frac{\pi}{2}\right]\right\}. \tag{7b}$$

From equation (33.7) we see that they behave like spherical outgoing and incoming waves, respectively. This suggests forming linear combinations which behave like standing waves, *i.e.* like sine and cosine functions. These functions are known as spherical Bessel functions and spherical Neumann functions (they are also called spherical Bessel functions of the first and second kinds) and are respectively defined by

$$j_l(x)=\frac{1}{2}\left[h_l^{(1)}(x)+h_l^{(2)}(x)\right], \tag{8a}$$

$$n_l(x)=\frac{1}{2i}\left[h_l^{(1)}(x)-h_l^{(2)}(x)\right]. \tag{8b}$$

It follows from equation (6) that these functions are real for real values of x and l. From equations (7) one obtains the asymptotic behaviour, for large x:

$$j_l(x)\simeq\frac{1}{x}\cos\left[x-(l+1)\frac{\pi}{2}\right], \tag{9a}$$

$$n_l(x)\simeq\frac{1}{x}\sin\left[x-(l+1)\frac{\pi}{2}\right]. \tag{9b}$$

For small x (*i.e.* as $x\rightarrow 0$), the spherical Bessel and Neumann functions behave like

$$j_l(x)\simeq\frac{x^l}{1\cdot 3\ldots(2l+1)}, \tag{10a}$$

$$n_l(x)\simeq\frac{(-1)1\cdot 3\ldots(2l-1)}{x^{l+1}}. \tag{10b}$$

From equation (10a) we see that the spherical Bessel function is that solution of the radial wave equation (4) which is regular at the origin; all other solutions have a singularity there.

The general solution of the free particle equation

$$(\nabla^2 + k^2)\psi = 0$$

can now be written as a superposition

$$\sum_{l=0}^{\infty} [A_l h_l^{(1)}(kr) + B_l h_l^{(2)}(kr)] P_l(\cos\theta). \tag{11}$$

In particular, we saw in section 32 that it possesses a plane wave solution

$$\psi_0 = e^{ikz} = e^{ikr\cos\theta} \tag{12}$$

which is regular everywhere. Hence it will contain only Bessel functions in its expansion which can be shown to be*

Plane wave expansion

$$\psi_0 = e^{ikz} = \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos\theta). \tag{13}$$

We can decompose ψ_0 into outgoing and incoming waves

$$\psi_0 = \psi_0^{(1)} + \psi_0^{(2)} \tag{14a}$$

where

$$\psi_0^{(s)} = \sum_{l=0}^{\infty} (2l+1) i^l \tfrac{1}{2} h_l^{(s)}(kr) P_l(\cos\theta). \tag{14b}$$

We return to the scattering equation (3). Outside the range of the potential, *i.e.* for $r > a$, this equation is identical with the free particle equation, so that for $r > a$, the general solution is still given by equation (11). For comparison with the plane wave expansion, equations (13) and (14), it will be convenient to write the general solution, equation (11), as

$$\psi = \sum_{l=0}^{\infty} (2l+1) i^l \tfrac{1}{2} [A_l h_l^{(1)}(kr) + B_l h_l^{(2)}(kr)] P_l(\cos\theta), \quad (r > a). \tag{11a}$$

Inside the range of the potential, *i.e.* for $r < a$, the scattering equation (3) and hence its solutions differ from the free particle case. One can show generally that for a given value of l this equation possesses two linearly independent solutions which can be chosen to behave like r^l and r^{-l-1} respectively at the origin†. They are the generalizations of the two solutions $j_l(kr)$ and $n_l(kr)$ [*cf* equations (10)] in the free particle case.

---

* *cf* Mott and Massey[21].

† This is the case provided $U(r)$ is not more singular than $1/r$ at the origin. We assume the potential to satisfy this condition.

Only the bounded solutions are physically admissible. For $l>0$ this follows from the fact that

$$\int_V |\psi|^2 d\mathbf{r},$$

representing the probability of finding the particle in a volume V, must be finite. For a small region V surrounding the origin, where the irregular solution behaves like r^{-l-1}, this integral diverges if $l>0$.

We shall now prove that for $l=0$ too, only the regular bounded solution is admissible, *i.e.* $(r\psi)_{r=0}$ must be zero (*cf* Dirac[15], pp. 155–156). We have

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{df(r)}{dr}\right)=\frac{1}{r}\frac{d^2}{dr^2}\{rf(r)\}, \tag{15}$$

and it is the singular behaviour of these expressions at $r=0$ which causes the trouble. We can write the radial equation (3) for $l=0$ as

$$\frac{1}{r}\frac{d^2}{dr^2}\{rG_0(r)\}+[k^2-U(r)]G_0(r)=0. \tag{16}$$

Consider first the special case of a free particle of zero energy:

$$k^2=U=0.$$

Equation (16) then becomes

$$\frac{1}{r}\frac{d^2}{dr^2}\{rG_0(r)\}=0 \tag{17}$$

and has the irregular solution

$$G_0(r)=\frac{1}{r}. \tag{18}$$

The general scattering equation (32.6):

$$[\nabla^2+k^2-U(r)]\psi=0, \tag{19}$$

in our special case with $k^2=U=0$ and $\psi=G_0(r)$ reads

$$\nabla^2 G_0(r)=0. \tag{20}$$

But the solution (18) of equation (17) is *not* a solution of the specialized scattering equation (20), since

$$\nabla^2\frac{1}{r}=-4\pi\delta(\mathbf{r});$$

i.e. it is a solution everywhere except at the origin $r=0$. Similarly in the general case $k^2\neq 0\neq U(r)$, if equation (16) possesses a solution

$G_0(r)$ which behaves like $1/r$ at the origin, this is not a solution of the scattering equation (19). For

$$\begin{aligned}\nabla^2 G_0(r) + [k^2 - U(r)]G_0(r) \\ &= \nabla^2 G_0 - \frac{1}{r}\frac{d^2}{dr^2}(rG_0) + \left\{\frac{1}{r}\frac{d^2}{dr^2}(rG_0) + [k^2 - U]G_0\right\} \\ &= \nabla^2 G_0 - \frac{1}{r}\frac{d^2}{dr^2}(rG_0),\end{aligned}$$

and if G_0 behaves like $1/r$ near the origin, this is equal to

$$\nabla^2 \frac{1}{r} = -4\pi\delta(\mathbf{r})$$

near the origin and hence is not a solution of the scattering equation at the origin. This shows that we must always consider only those solutions of the radial scattering equation (3) for which $r\psi \to 0$ as $r \to 0$, *i.e.* the bounded solutions.

From the required asymptotic form (32.9) it follows that $\psi - e^{ikz}$ must contain only outgoing waves. From equations (14) and (11a) this implies

$$B_l = 1. \tag{21}$$

The physical interpretation of this is of course that the incoming part of the plane wave which originates at large distances from the origin is not affected by the scattering centre. This only affects the outgoing waves, distorting them from what they would have been in the free particle case.

We can now rewrite equation (11a) using equations (21) and (8) as

$$\begin{aligned}\psi &= \sum_{l=0}^{\infty}(2l+1)i^l \tfrac{1}{2}[A_l h_l^{(1)}(kr) + h_l^{(2)}(kr)]P_l(\cos\theta), \quad (r>a), \\ &= \sum_{l=0}^{\infty}(2l+1)i^l \tfrac{1}{2}[(A_l+1)j_l(kr) + i(A_l-1)n_l(kr)]P_l(\cos\theta), \quad (r>a). \quad (22)\end{aligned}$$

As we have seen above, ψ is to be a regular solution. For each l-value, the regular solution $G_l(r)$ of the scattering equation (3) is completely determined apart from a constant multiplicative factor. It behaves like r^l at the origin. By a suitable choice of the multiplicative factor, we can make the first non-vanishing derivative at the origin, $(d^lG_l(r)/dr^l)_{r=0}$, real. Then $G_l(r)$ is real for all values of r, since the scattering equation (3) is a real differential equation, *i.e.* l, $U(r)$ and k^2

are real. Since ψ, equation (22), is a regular solution, it follows that, for *all* values of $r>a$,

$$(A_l+1)\left[j_l(kr)+i\frac{A_l-1}{A_l+1}n_l(kr)\right]=C_lG_l(r),\quad (r>a), \tag{23}$$

where C_l is a complex number determined by the normalization of the real function $G_l(r)$. Consequently the ratio

$$i\frac{A_l-1}{A_l+1} \tag{24}$$

is completely determined and is real since $j_l(kr)$ and $n_l(kr)$ are real. It follows at once that A_l is a phase factor of modulus unity which we can write

$$A_l=\exp(2i\delta_l), \tag{25}$$

where δ_l is real. Thus we finally obtain for the wave function which represents the solution of our scattering problem

$$\psi=\sum_{l=0}^{\infty}(2l+1)i^l\tfrac{1}{2}[\exp(2i\delta_l)h_l^{(1)}(kr)+h_l^{(2)}(kr)]P_l(\cos\theta),\quad (r>a). \tag{26}$$

Equation (25) allows a very simple physical interpretation. Each partial wave with a given l-value corresponds to particles of angular momentum l. We are treating elastic scattering, so that there are as many outgoing particles of angular momentum l as are incident on the scattering centre with angular momentum l. Hence the normalizations of the outgoing and incoming parts of each partial wave must be the same. From equation (22) this at once leads to $|A_l|^2=1$. Only a phase factor is left undetermined. Hence, compared to the free particle case, each outgoing partial wave has experienced a *phase-shift* δ_l. As a result the coherent sum of the partial waves no longer gives a plane but a distorted wave, corresponding to the scattering that has taken place.

From equations (26) and (14) one finds the scattered wave

$$\begin{aligned}\psi_{sc}&=\psi-\psi_0\\&=\sum_{l=0}^{\infty}(2l+1)i^l\tfrac{1}{2}h_l^{(1)}(kr)[\exp(2i\delta_l)-1]P_l(\cos\theta),\quad (r>a),\end{aligned} \tag{27}$$

which has the asymptotic form, for large values of r,

$$\psi_{sc}\simeq\frac{e^{ikr}}{2ikr}\sum_{l=0}^{\infty}(2l+1)[\exp(2i\delta_l)-1]P_l(\cos\theta). \tag{28}$$

Comparison with equation (32.9) gives the scattering amplitude

$$\left.\begin{aligned} f(\theta) &= \frac{1}{2ik}\sum_{l=0}^{\infty}(2l+1)[\exp(2i\delta_l)-1]P_l(\cos\theta) \\ &= \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)\exp(i\delta_l)\sin\delta_l P_l(\cos\theta). \end{aligned}\right\} \tag{29}$$

Equations (29) represent the basic result of the partial wave analysis. The scattering amplitude (and hence the differential and total cross-sections) are expressed in terms of the phase shifts. These form an infinite set. We shall see presently that often only a few phases make a significant contribution to the cross-section, *i.e.* the remaining phase shifts δ_l are very small. It is in these cases that the partial wave analysis is so useful. The differential cross-section is now given by equations (29) as a finite polynomial in $\cos\theta$ whose coefficients are known functions of the phase shifts. By fitting these expressions to the experimentally determined differential cross-sections one can try to determine the phase shifts.

From equations (29), (32.11) and (32.12) one easily obtains the total cross-section, using the orthonormality of the Legendre polynomials [equations (13.16) and (13.17)]. One finds

$$\sigma_{\text{tot}}(k) = \frac{4\pi}{k^2}\sum_{l=0}^{\infty}(2l+1)\sin^2\delta_l(k), \tag{30}$$

where we have written $\delta_l(k)$ and $\sigma_{\text{tot}}(k)$ to stress the energy dependence of phase shifts and cross-sections. One easily proves the *optical theorem* from equations (29) and (30) which relates the imaginary part of the forward scattering amplitude (*i.e.* $\theta=0$) to the total cross-section [using the fact that $P_l(1)=1$]:

$$\sigma_{\text{tot}} = \frac{4\pi}{k}\operatorname{Im} f(0), \tag{31}$$

(*cf* Exercises VIII.6–7).

To find the phases for a given potential is in general a difficult problem. For a few potentials one can find the phases from an exact solution of the scattering problem. In general, one has to solve the scattering problem by a numerical integration of the radial wave equation, going outwards from the centre of the potential, matching the logarithmic derivative of the wave function at the surface of the potential (*cf* Exercise III.5), and proceeding to large distances to find the asymptotic behaviour and hence the phase shifts. Under certain circumstances various special methods exist for determining the phases. These matters are fully discussed in the references given at the beginning of this section. Here we only wish to discuss a few points.

We stated above that, in general, only a few partial waves will be of importance. Classically, one can easily see this. A particle of momentum $\hbar k$ and angular momentum $\hbar l$ will have an impact parameter b defined by

$$\hbar l = \hbar k b \tag{32}$$

and this particle will only be scattered if the particle comes within the range of the potential, *i.e.* if the impact parameter b is less than the range of the potential $a : b < a$. Hence we get the condition, from equation (32),

$$l < ka \tag{33}$$

for the lth partial wave to be scattered.

A similar result can be derived quantum-mechanically, but it is not easy to give a rigorous proof of general validity. We shall only give a plausibility argument. The radial equation for the lth partial wave of a free particle, equation (4), contains the term $-l(l+1)/r^2$. This term represents a repulsive potential, of increasing strength as l increases. We can interpret it as a centrifugal barrier which prevents a particle of high angular momentum from penetrating to small distances r. This interpretation is justified by the fact that the solution $j_l(kr)$ of equation (4) is steeply increasing for small (kr). From equation (10a) it is proportional to $(kr)^l$ and one can show that it has its first maximum near $kr \simeq l$. Hence there is only a very small probability of particles penetrating to distances

$$r < r_l \equiv l/k. \tag{34}$$

If we again consider scattering of particles of a given momentum $\hbar k$ by a potential of definite range a, it follows that little or no scattering will occur for partial waves of sufficiently high l-value. The centrifugal barrier prevents these particles from penetrating into the range of the potential: $a < r_l$. Thus scattering will be negligible for partial waves for which $l > ka$, agreeing with the classical criterion, equation (33).

We can think of this criterion as follows. For a given potential, we fix our attention on a definite l-wave. At sufficiently low energies, this partial wave will in general not experience appreciable scattering. As the energy is increased, *i.e.* as k is increased sufficiently, the scattering of the partial wave gradually becomes important. The region where it is no longer negligible can be estimated from equation (33). At sufficiently low energies, only s-wave scattering $(l=0)$ will occur. As the energy is increased, the p-wave contribution to the scattering will become important and so on.

From equations (29) we see that the angular distribution undergoes a definite change as successive partial waves come into play: if only partial waves with l-values less or equal to L contribute significantly to the scattering, the angular distribution will be given by a polynomial of degree not greater than $2L$ in $\cos\theta$. In particular it is isotropic if there is only s-wave scattering.

We shall next derive an exact expression for the phase shifts in terms of the wave function. Defining

$$u_l(kr) = krj_l(kr), \tag{35}$$

equation (4) becomes

$$\frac{d^2u_l}{dr^2} + \left[k^2 - \frac{l(l+1)}{r^2}\right]u_l = 0, \tag{36}$$

and from equation (9a), u_l behaves asymptotically like

$$u_l(kr) \simeq \sin\,(kr - \tfrac{1}{2}l\pi). \tag{37}$$

Similarly writing

$$v_l(r) = krG_l(r),$$

equation (3) becomes

$$\frac{d^2v_l}{dr^2} + \left[k^2 - U(r) - \frac{l(l+1)}{r^2}\right]v_l = 0. \tag{38}$$

From equations (23), (25) and (9), we see that v_l behaves asymptotically like

$$v_l(r) \simeq \sin\,(kr - \tfrac{1}{2}l\pi + \delta_l). \tag{39}$$

From equations (36) and (38) one obtains

$$v_l\frac{d^2u_l}{dr^2} - u_l\frac{d^2v_l}{dr^2} = -U(r)u_lv_l. \tag{40}$$

We integrate this equation from $r=0$ to r. The contribution to the left-hand side from the lower limit vanishes, since $u_l(0) = v_l(0) = 0$. [$u_l(0) = 0$ follows directly from equations (35) and (10a). $v_l(0) = 0$ follows since $G_l(r)$ is to be the bounded solution of the scattering equation.] Hence

$$v_l(r)\frac{d}{dr}u_l(kr) - u_l(kr)\frac{d}{dr}v_l(r) = -\int_0^r U(r')u_l(kr')v_l(r')dr'. \tag{41}$$

If we now let r tend to infinity and substitute on the left-hand side the asymptotic forms (37) and (39), we obtain

$$k\sin\delta_l = -\int_0^\infty U(r)u_l(kr)v_l(r)dr. \tag{42}$$

In the energy region where a given δ_l is small, due to the centrifugal barrier as discussed above, we can calculate δ_l in first order perturbation theory, *i.e.* use the unperturbed wave function in equation (42).

Using equations (32.4) and (32.5) and writing δ_l for $\sin\delta_l$, we find

$$\delta_l = -\frac{\pi\mu}{\hbar^2}\int_0^\infty V(r)[J_{l+\frac{1}{2}}(kr)]^2 r\,dr. \tag{43}$$

This equation shows that, in the conditions under which it is valid, an attractive potential (*i.e.* $V(r) \leqslant 0$ for *all* r) leads to positive phases, a repulsive potential (*i.e.* $V(r) \geqslant 0$ for *all* r) to negative phases. This result can be proved generally (*cf* Exercise **VIII**.8).

Using equation (43) for all partial waves and assuming the δ_l small is equivalent to using Born approximation. To see this, we substitute it into equation (29), replacing $\exp(2i\delta_l) - 1$ by $2i\delta_l$ and use the identity [*cf* Watson, *loc. cit.*, p. 363, eq. (3)]:

$$\frac{\sin Kr}{Kr} = \sum_{l=0}^{\infty}(2l+1)P_l(\cos\theta)[j_l(kr)]^2, \tag{44}$$

where $K = 2k\sin(\theta/2)$ as in equation (34.5). In this way one again obtains the Born approximation (34.3) for the scattering amplitude.

The above considerations must be modified if one deals with particles with spin. The detailed analysis becomes considerably more complicated although the basic ideas are not changed. One may now no longer simply consider states of given orbital angular momentum l. Instead one must decompose the wave function into states of definite total angular momentum and consider these separately. This is, for example, what has to be done in nucleon–nucleon scattering.

EXERCISES VIII

VIII.1. Verify that equation (33.7) is an exact solution of equation (33.6).

VIII.2. Obtain, for the inelastic scattering process considered in section 36, the transition probability per unit time to a state in which the scatterer is in the state $\varphi_s(q)$ and the scattered particle has a momentum vector in the interval $(\hbar\mathbf{k}_s, \hbar\mathbf{k}_s + \hbar d\mathbf{k}_s)$.

VIII.3. Obtain the differential cross-section for the inelastic scattering process considered in section 36, in first Born approximation.

VIII.4. Obtain the differential scattering cross-section, equation (36.16) for the elastic scattering of two identical particles.

? **VIII.5.** Calculate the differential and total cross-sections for elastic scattering of fast electrons by atomic hydrogen (*cf* Exercise **III**.6), neglecting exchange effects due to the identity of incoming and atomic electrons.

VIII.6. Derive equation (38.30).

VIII.7. Derive the optical theorem, equation (38.31).

VIII.8. Prove that a repulsive potential leads to negative phase shifts δ_l, an attractive potential to positive phase shifts.

VIII.9. Obtain an expression for the phase shift δ_0 for *s*-wave scattering by a spherical square-well potential: $V(r) = -V_0$ if $r < a$; $V(r) = 0$ if $r > a$; $(V_0 > 0)$. Show that, though $\delta_0 \rightarrow 0$ as the bombarding energy tends to zero, the cross-section tends, in general, to a finite limit.

IX

INTRODUCTION TO GROUP-THEORETICAL IDEAS

In earlier chapters of this book, we have repeatedly dealt with angular momentum, parity and symmetry properties of wave functions. These play a fundamental role in many problems, particularly in the case of complicated systems where exact, quantitative calculations are difficult and where it is important to gain as much information as possible about the system by general methods.

We shall now consider these symmetry properties more systematically. For this purpose, we shall introduce the basic ideas of group theory. These are exceedingly simple, and, as we shall see, have a very beautiful unifying effect, allowing many apparently diverse questions to be answered by essentially the same methods. These methods have been developed into an elaborate mathematical theory enabling one to solve very complicated problems. We shall of course restrict ourselves to the fundamental concepts underlying these methods, which should not only serve as an introduction to more specialized books on group theory* but should enable the reader to treat many important problems in atomic and nuclear physics.

39. Groups of transformations

Consider a system of N identical particles of mass m whose Hamiltonian referred to a given system of rectangular coordinates C is

$$H = \sum_i \frac{1}{2m} \mathbf{p}_i^2 + \sum_{i<j} V(|\mathbf{r}_i - \mathbf{r}_j|). \tag{1}$$

If we consider transformations to other coordinate systems, we find that some of these leave the Hamiltonian (1) unchanged. We shall say that H is *invariant* or *symmetric* under these transformations. These particular transformations include:

(*I*) Permutations of the N identical particles;

(*II*) Reflection of the coordinate system C in the origin, *i.e.* the transformation $\mathbf{r} \to -\mathbf{r}$, going from a right-handed to a left-handed coordinate system and vice versa;

(*III*) Spatial rotations of the coordinate system about the fixed origin;

(*IV*) Rotations of the coordinate system about an axis fixed in space. (We shall usually take this to be the z-axis.)

---

* A particularly lucid and simple treatment will be found in the book by van der Waerden[23].

The invariance of (1) with respect to permutations of particles follows from the identity of the particles; permuting them simply implies a rearrangement of terms in (1). The invariance of (1) with respect to the three other sets of transformations follows from the fact that (1) only depends on the magnitude of certain vectors, but not on their sense or orientation.

Applying a homogeneous magnetic field in the z-direction to the atom, removes the spherical symmetry (*III*), but the system will still be axially symmetric about the z-axis, like the magnetic interaction term, equation (31.25), which must now be added to the Hamiltonian (1)*.

If we supplement each of the sets of transformations (*I*) to (*IV*) by the identity transformation, which leaves the coordinate system unchanged, then each set of coordinate transformations forms a *group*, *i.e.* it has these properties: (*i*) two transformations of the set, applied successively, are equivalent to a single transformation of the set; (*ii*) corresponding to each transformation $C \to C'$ of the set, there exists an inverse transformation, $C' \to C$, belonging to the set.

To define a group of coordinate transformations generally, we denote the transformation $C \to C'$ by t and the coordinates in these two systems by q and q' (standing for $q_1, q_2, \ldots$ and $q_1', q_2', \ldots$), so that

$$t : q \to q' = tq \tag{2}$$

(*i.e.* t operating on q transforms it into $q' = tq$). We can define a group of coordinate transformations as follows.

A set of transformations $\{s, t, u, \ldots\}$ forms a group, if

(*i*) any two transformations s and t applied successively are a transformation of the set; we write this transformation ts and call it the product transformation; *i.e.* if $s : q \to q' = sq$, and $t : q' \to q'' = tq'$, then $(ts) : q \to q'' = tsq$;

(*ii*) to every transformation t of the set, there exists an inverse transformation, denoted by t^{-1}, such that if $t : q \to q' = tq$, then $t^{-1} : q' \to q = t^{-1}q'$.

It follows from (*ii*) that the unit element or identity e which leaves the coordinate system unchanged, $e : q \to q$, has the properties that $tt^{-1} = t^{-1}t = e$, and $te = et = t$.

Our notation tq and terminology 'product', 'unit element', *etc* must not be interpreted too naively as multiplication *etc* as t stands for a quite general transformation. Consider the group $\mathfrak{R}_2$ of axial rotations about the z-axis, example (*IV*) above ($\mathfrak{R}_2$ as we are dealing with two-dimensional rotations). A rotation of the coordinate system through an angle α (we shall denote the corresponding transformation by t_α) only changes the azimuthal angle φ of the polar coordinates

---

* For the moment we are neglecting spin, which we shall only introduce in section 42.

(r, θ, φ) of a point $P; t_\alpha : \varphi \rightarrow \varphi' = t_\alpha \varphi = \varphi - \alpha$, as can be seen from *Figure 14*. One easily sees that $(t_\alpha)^{-1} = t_{-\alpha}$, that the identity transformation is $e = t_0$ (*i.e.* $\alpha = 0$, no rotation) and the product of t_α and t_β has the property

$$t_\alpha t_\beta = t_{\alpha+\beta} = t_{\beta+\alpha} = t_\beta t_\alpha. \tag{3}$$

This last equation states the geometrically obvious fact that two successive rotations about the same axis carried out in either order lead to the same final configuration. We say the transformations of the axial rotation group commute; it is called an *Abelian* group. The reflection or inversion group $\mathfrak{G}_I$ [example (*II*) above], consisting of only the two transformations $e : \mathbf{r} \rightarrow \mathbf{r}$ and $i : \mathbf{r} \rightarrow -\mathbf{r}$ is also Abelian.

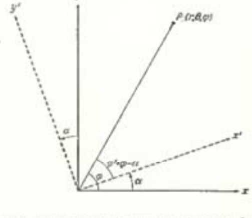

Figure 14. Axial rotation of coordinate system about z-axis through angle α

The group of spatial rotations $\mathfrak{R}_3$, or as we shall say more briefly, the rotation group $\mathfrak{R}_3$ [example (*III*) above], however, is non-commutative: two finite rotations, carried out in either order, in general lead to two different configurations. Similarly, the permutation group $\mathfrak{S}_N$ of N objects [example (*I*) above] is non-commutative for $N > 2$.

We now consider the effect of the transformation t, equation (2), on a function $f(q)$. It is equivalent to a substitution of new coordinates:

$$f(q) = f(t^{-1}q') = f'(q'). \tag{4}$$

Given the transformation t, equation (4) defines a new function $f'(q)$ for any given function $f(q)$, *i.e.* it defines an operator T:

$$T : f(q) \rightarrow f'(q) = Tf(q) = f(t^{-1}q). \tag{5}$$

The operator T transforms the function $f(q)$ into that new function $f'(q)$ which has the same value at the same point referred to the new coordinates. To give a simple example, suppose $f(q) = \sin q$, and $q' = tq = q^2$, so that $t^{-1}q = \sqrt{q}$. Then $f'(q) = f(t^{-1}q) = \sin \sqrt{q}$.

For the applications of group theory to quantum mechanics, it is of fundamental importance that if the set of coordinate transformations $\{s, t, u, \dots\}$ forms a group, so does the set of operators $\{S, T, U, \dots\}$. This is easily proved. Suppose $q' = sq$, $q'' = tq'$ and $q'' = uq$, so that $u = ts$. We have, from equation (5), that $Sf(q') = f(q)$, and putting $\varphi(q) = Sf(q)$, that $T\varphi(q'') = \varphi(q')$, *i.e.* $TSf(q'') = f(q)$. But we also have $Uf(q'') = f(q)$, and, since $f(q)$ is an arbitrary function, it follows that $U = TS$. One similarly proves that if s^{-1} is the coordinate transformation $q = s^{-1}q'$, then the operator associated with this transformation is $S^{-1}: f(q) \to f(q') = S^{-1}f(q)$, *i.e.* the inverse operator to S. Hence the operators $\{S, T, U, \dots\}$ have exactly the same laws of combination as the coordinate transformations $\{s, t, u, \dots\}$, and hence form a group which differs in no essentials from the group $\{s, t, u, \dots\}$ [*].

The functions $f(q)$ that occur in quantum mechanics are of two kinds. $f(q)$ may be an operator (in which case the transformation t must of course also be applied to any differential operators, such as $-i\hbar\partial/\partial q$, that may occur). Alternatively, it may be a wave function describing the system under consideration, so that T becomes an operator in the corresponding function space. We can then think of the coordinate transformation t in configuration space as inducing a transformation in the function space, by means of the operator T.

The operators T occurring in quantum mechanics are always unitary. For the transformation groups that occur, in particular the four examples we gave, leave the inner product invariant. They correspond simply to a change of the variables of integration. Consider a spatial rotation $t: \mathbf{r} \to \mathbf{r}' = t\mathbf{r}$. Since $d\mathbf{r}' = d\mathbf{r}$, it follows from equation (4) that for any N-particle wave functions f and g,

$$\begin{aligned}(g, f) &= \int d\mathbf{r}_1 \dots d\mathbf{r}_N\, \overline{g(\mathbf{r}_1, \dots \mathbf{r}_N)}\, f(\mathbf{r}_1, \dots \mathbf{r}_N)\\ &= \int d\mathbf{r}_1' \dots d\mathbf{r}_N'\, \overline{g'(\mathbf{r}_1', \dots \mathbf{r}_N')}\, f'(\mathbf{r}_1', \dots \mathbf{r}_N')\\ &= (g', f') = (Tg, Tf), \qquad (6)\end{aligned}$$

i.e. the operator T is unitary [*cf* equation (18.15)]. In the same way one proves this result for other transformation groups.

40. Rotation operators

We consider the spatial rotations of a system of N particles, with position coordinates (x_i, y_i, z_i), $i = 1, 2, \dots N$. If t_z denotes a coordinate transformation rotating the coordinate system through an angle $\delta\alpha$ about the z-axis, then

$$x_i' = t_z x_i = x_i + \delta x_i, \quad (i = 1, 2, \dots N), \textit{ etc}$$

where, as one easily sees by going over to spherical polar coordinates,

$$\delta x_i = y_i \delta\alpha, \quad \delta y_i = -x_i \delta\alpha, \quad \delta z_i = 0.$$

---

[*] In fact, the groups $\{s, t, u, \dots\}$ and $\{S, T, U, \dots\}$ are *isomorphic*, *i.e.* to every element s corresponds exactly one element S and vice versa, and if $st = u$, then $ST = U$. Thus the group properties of two isomorphic groups are the same. We have merely labelled the elements differently.

Writing $\varphi(x_i, y_i, z_i)$ for $\varphi(x_1, \ldots z_N)$, one finds, on account of equations (39.4–5), that

$$\begin{aligned}
\varphi'(x_i, y_i, z_i) &= T_z\varphi(x_i, y_i, z_i) = \varphi(t_z^{-1}x_i, t_z^{-1}y_i, t_z^{-1}z_i) \\
&= \varphi(x_i - \delta x_i, y_i - \delta y_i, z_i - \delta z_i) \\
&= \varphi(x_i, y_i, z_i) - \sum_{i=1}^{N}\left(\frac{\partial\varphi}{\partial x_i}\delta x_i + \ldots + \ldots\right) \\
&= \varphi(x_i, y_i, z_i) - \delta\alpha\sum_{i=1}^{N}\left(y_i\frac{\partial}{\partial x_i} - x_i\frac{\partial}{\partial y_i}\right)\varphi(x_i, y_i, z_i) \\
&= \left(1 - \frac{\delta\alpha}{i\hbar}L_z\right)\varphi(x_i, y_i, z_i),
\end{aligned}$$

so that

$$T_z = \left(1 - \frac{\delta\alpha}{i\hbar}L_z\right), \tag{1}$$

where L_z is the z-component of the total orbital angular momentum operator of the system (Exercise II.1). Since analogous expressions hold for the operators T_x and T_y, corresponding to infinitesimal rotations through an angle $\delta\alpha$ about the x and y axes, it follows that the angular momentum operators completely determine the infinitesimal rotation operators T_x, T_y, T_z. But one can further prove that the transformation properties of a wave function under a finite rotation are completely determined by its behaviour under the infinitesimal rotations. For since the rotations form a group, the operator of a finite rotation can be built from the infinitesimal rotation operators by carrying out successive infinitesimal rotations, *i.e.* by a process of integration. Hence the angular momentum properties of a system (*i.e.* the effect of the operators L_x, L_y, L_z on its wave function) completely determine its properties under spatial rotations.

41. Invariants of groups of transformations

A function $f(q)$ (which again may be a wave function or an operator) is invariant under a group of transformations $\{s, t, u, \ldots\}$ if for every transformation t of the group

$$f'(q) = Tf(q) = f(t^{-1}q) = f(q). \tag{1}$$

For example, the spatial rotation group $\mathfrak{R}_3$ leaves the magnitudes of vectors, or functions depending only on the magnitudes of vectors, invariant. A very general Hamiltonian of an N-particle system, which is invariant under rotations, *i.e.* spherically symmetric, is

$$H_0 = -\sum_{i=1}^{N}\frac{\hbar^2}{2m_i}\nabla_i^2 + V(r_1, r_2, \ldots r_N; r_{12}, r_{13}, \ldots) \tag{2}$$

where the interaction potential V is an arbitrary function of the distances $r_i = |\mathbf{r}_i|$ and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$.

One can easily see that if an operator $\Gamma(q)$ is invariant under a transformation group $\{s, t, u, \dots\}$, then it commutes with every operator of the corresponding group $\{S, T, U, \dots\}$:

$$[\Gamma(q), S]=0, \quad etc. \tag{3}$$

For it follows from (1) for any wave function $\varphi(q)$, that

$$S\Gamma(q)\varphi(q)=\Gamma'(q)\varphi'(q)=\Gamma(q)\varphi'(q)=\Gamma(q)S\varphi(q),$$

i.e. (3) holds. Conversely, if $\varphi(q)$ is an arbitrary function and we define $\psi(q)=S^{-1}\varphi(q)$, then equation (3), which must in particular hold for $\psi(q)$ (since it holds generally), gives

$$\{\Gamma'(q)-\Gamma(q)\}\,\varphi(q)=0, \tag{4}$$

and since $\varphi(q)$ is an arbitrary function, equation (1) follows.

The results we have just obtained can be applied with advantage to the discussion of the constants of the motion (or the good quantum numbers, as it is often put) of a system. We give two examples.

(*1*) A spherically symmetric Hamiltonian H [*e.g.* equation (2)] is invariant under spatial rotations. Hence it commutes with all rotation operators, in particular with the infinitesimal rotation operators, equation (40.1), and hence with the total orbital angular momentum operators:

$$[H, L_x]=0, \quad etc. \tag{5}$$

Since L^2 is also spherically symmetric and $L^2=L_x^2+L_y^2+L_z^2$ we also obtain

$$[H, L^2]=0 \qquad [L^2, L_x]=0, \quad etc. \tag{6}$$

Hence, for a spherically symmetric Hamiltonian, H, L^2 and one of the components, say L_z, are constants of the motion that can be fixed. (Since the three components L_x, L_y, L_z do not commute with each other, we can only fix one of them.) We have here obtained one of the main results about angular momentum (already derived in Chapters II, IV and VI), with very little labour. Furthermore, the significance of equations (5), which in section 25 we introduced somewhat artificially, is now apparent.

(*2*) Consider the reflection or inversion group $\mathfrak{G}_i$. If I is the operator corresponding to the inversion of the coordinate system, $i:\mathbf{r}\to-\mathbf{r}$, then

$$I\psi=\pm\psi \tag{7}$$

for wave functions of even and odd parity respectively. If the Hamiltonian H of a system is invariant under the reflection group, then H commutes with I, and so parity is a good quantum number.

In applying the inversion transformation, $i:\mathbf{r}\to-\mathbf{r}$, one must carefully distinguish between vectors and pseudo-vectors which transform in the same way under rotations, but differently under reflection. If $\mathbf{a}(v)$ is a vector and $\mathbf{a}(pv)$ is a pseudo-vector, then

$$i\mathbf{a}(v)=-\mathbf{a}(v), \quad i\mathbf{a}(pv)=\mathbf{a}(pv). \tag{8}$$

It follows from the definition of inversion that such quantities as momentum, $\mathbf{p}=m\dot{\mathbf{r}}$, and force are vectors, whereas angular momentum, $\mathbf{l}=\mathbf{r}\wedge\mathbf{p}$, is a pseudo-vector:

$$i:\mathbf{l}=\mathbf{r}\wedge\mathbf{p}\rightarrow(-\mathbf{r})\wedge(-\mathbf{p})=\mathbf{l}. \tag{9}$$

42. Spin

We now generalize our methods to be applicable to particles with spin $\frac{1}{2}$. For simplicity we shall consider a single particle but the generalizations to several particles will be obvious.

As in section 26, we specify a particle of spin $\frac{1}{2}$ by a two-component wave function

$$\Phi(\mathbf{r}, s_z)=\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta \tag{1}$$

being a vector in the product space $\mathfrak{H}\times\mathfrak{S}$. The operators in this space will now generally be functions of $\mathbf{r}$ and the Pauli vector $\boldsymbol{\sigma}$, $\Gamma=\Gamma(\mathbf{r}, \boldsymbol{\sigma})$. We consider the effect of a group of transformations $\{s, t, u, \ldots\}$ on such operators Γ and wave functions Φ.

Applied to a wave function, a coordinate transformation t now induces transformations both in the orbital space $\mathfrak{H}$ and the spin space $\mathfrak{S}$. Though our treatment and results are generally valid, we shall consider the rotation group, in order to be able to interpret our results physically at every stage.

Consider the spatial rotation $t:\mathbf{r}\rightarrow\mathbf{r}'=t\mathbf{r}$ from the coordinate system C to C'. As a result the orbital wave functions $\varphi_i(\mathbf{r})$ become $\varphi_i(t^{-1}\mathbf{r}')=\varphi_i'(\mathbf{r}')$, as in section 39. But referred to the new coordinates the spin of the particle will point in a different direction, the spin wave functions α and β will be transformed into linear combinations of the new spin wave functions α' and β':

$$\alpha=c_1\alpha'+c_2\beta', \qquad \beta=c_3\alpha'+c_4\beta'. \tag{2}$$

We shall not derive this linear transformation, but merely state that for a given transformation t it is completely determined, $c_i=c_i(t)$, and that it is a unitary transformation.

Combining these results, we have that

$$\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta=\varphi_1(t^{-1}\mathbf{r}')\{c_1\alpha'+c_2\beta'\}+\varphi_2(t^{-1}\mathbf{r}')\{c_3\alpha'+c_4\beta'\}. \tag{3}$$

As in section 39, we can express this by saying that t induces a transformation T in the function space $\mathfrak{H}\times\mathfrak{S}$ defined by

$$T[\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta]=\varphi_1(t^{-1}\mathbf{r})\{c_1\alpha+c_2\beta\}+\varphi_2(t^{-1}\mathbf{r})\{c_3\alpha+c_4\beta\}. \tag{4a}$$

We can decompose the operator T into two operators T^o and T^s operating in orbital and spin space respectively, defined by

$$T^o\varphi(\mathbf{r})=\varphi(t^{-1}\mathbf{r}) \tag{4b}$$

$$T^s[\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta]=\varphi_1(\mathbf{r})\{c_1\alpha+c_2\beta\}+\varphi_2(\mathbf{r})\{c_3\alpha+c_4\beta\}, \tag{4c}$$

and one easily shows that

$$T=T^oT^s=T^sT^o. \tag{5}$$

The permutation operators, P_{12}, P_{12}^x and P_{12}^σ which we introduced in section 28, are, of course, particular cases of T, T^x and T^σ corresponding to the symmetric group $\mathfrak{S}_2$. The operators T^x are of course identical with the operators T which we had in the spinless theory (section 41).

We can now apply the operators T, T^x and T^σ to an operator $\Gamma(\mathbf{r}, \boldsymbol{\sigma})$ and obtain

$$T\Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(t^{-1}\mathbf{r}, t^{-1}\boldsymbol{\sigma}), \quad T^x\Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(t^{-1}\mathbf{r}, \boldsymbol{\sigma}),$$
$$T^\sigma\Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(\mathbf{r}, t^{-1}\boldsymbol{\sigma}). \tag{6}$$

Analogously to section 41, we can now define the invariance of the operator $\Gamma(\mathbf{r}, \boldsymbol{\sigma})$ under the group $\{s, t, \ldots\}$, considered to induce transformations in the total, orbital or spin spaces respectively, by requiring that for all transformations t of the group:

$$T\Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(\mathbf{r}, \boldsymbol{\sigma}), \quad T^x\Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(\mathbf{r}, \boldsymbol{\sigma}) \quad T^\sigma\Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(\mathbf{r}, \boldsymbol{\sigma}), \tag{7}$$

and, as in section 41, one easily proves that these three equations respectively imply

$$[T, \Gamma(\mathbf{r}, \boldsymbol{\sigma})] = 0, \quad [T^x, \Gamma(\mathbf{r}, \boldsymbol{\sigma})] = 0, \quad [T^\sigma, \Gamma(\mathbf{r}, \boldsymbol{\sigma})] = 0 \tag{8}$$

for all operators of the groups $\{S, T, \ldots\}$, $\{S^x, T^x, \ldots\}$ and $\{S^\sigma, T^\sigma, \ldots\}$ respectively, and vice versa.

All the above considerations are easily generalized to systems of particles. An operator $\Gamma(\mathbf{r}, \boldsymbol{\sigma})$ becomes replaced by $\Gamma(\mathbf{r}_1, \mathbf{r}_2, \ldots; \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \ldots)$ and a wave function $\Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots; s_z^{(1)}, s_z^{(2)} \ldots)$ takes the place of $\Phi(\mathbf{r}, s_z)$.

We consider the spatial rotation group $\mathfrak{R}_3$. If t is such a rotation in configuration space, we shall call T^x, T^σ and T rotations in orbital, spin and total function space. In section 40 we obtained the infinitesimal rotation operators in orbital space:

$$T_x^x = \left(1 - \frac{\delta a}{i\hbar} L_x\right), \quad etc \tag{9a}$$

[equation (40.1)], **L** being the total orbital angular momentum vector. For infinitesimal rotations in spin space one expects by analogy (and one can prove this to be correct) that

$$T_x^\sigma = \left(1 - \frac{\delta a}{i\hbar} S_x\right), \quad etc \tag{9b}$$

S being the total spin vector of the system,

$$\mathbf{S} = \frac{\hbar}{2} \sum_i \boldsymbol{\sigma}(i). \tag{10}$$

From equations (5), (9a) and (9b) one obtains for the infinitesimal rotation operators in the total space

$$T_x = \left(1 - \frac{\delta\alpha}{i\hbar}\mathcal{J}_x\right), \quad \textit{etc} \tag{9c}$$

where $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the total angular momentum of the system.

We next consider the inversion group $\mathfrak{G}_I$. The inversion $i : \mathbf{r} \rightarrow \mathbf{r}' = -\mathbf{r}$ again induces operations I^x, I^σ and $I = I^x I^\sigma = I^\sigma I^x$ in the function spaces $\mathfrak{H}$, $\mathfrak{S}$ and $\mathfrak{H} \times \mathfrak{S}$. Since the Pauli vector $\boldsymbol{\sigma}$ represents an angular momentum and is therefore a pseudo-vector [*cf* equation (41.9)], inversion leaves it invariant, so that

$$I^\sigma \Gamma(\mathbf{r}, \boldsymbol{\sigma}) = \Gamma(\mathbf{r}, i\boldsymbol{\sigma}) = \Gamma(\mathbf{r}, \boldsymbol{\sigma}). \tag{11}$$

One easily sees that

$$\alpha' = I^\sigma \alpha = \alpha, \quad \beta' = I^\sigma \beta = \beta, \tag{12}$$

in other words, inversion does not affect spin space at all, I^σ is simply the unit operator and $I = I^x$. For if (x, y, z) is a coordinate system and (x', y', z') the inverted system, so that $x' = -x$, *etc*, then $\sigma_{z'} = \sigma_z$ since $\boldsymbol{\sigma}$ is a pseudo-vector. Hence, $\sigma_z \alpha = \alpha$ implies $\sigma_{z'} \alpha = \alpha$. But $\sigma_{z'} \alpha' = \alpha'$, and we know that the eigenfunctions of σ_z are non-degenerate, so that $\alpha' = \alpha$. It similarly follows that $\beta' = \beta$, proving equation (12).

We illustrate these results for the N-particle Hamiltonian

$$H_1 = H_0 + \sum_{i<j} V_{ij}(r_1, \ldots r_N; r_{12}, \ldots r_{N-1,N}) \boldsymbol{\sigma}_i \boldsymbol{\sigma}_j \tag{13}$$

where H_0 is the spin-independent Hamiltonian defined in equation (41.2) and the second term represents a spin-dependent interaction. The functions V_{ij} again are quite general functions of the arguments $r_1, \ldots, r_{12}, \ldots$ We see that both parts of the Hamiltonian are invariant under separate rotations in orbital and spin space, since $\boldsymbol{\sigma}_i \boldsymbol{\sigma}_j$ being a scalar product of two vectors is invariant under rotations, and of course, also under joint rotations. Furthermore, this Hamiltonian is also invariant under inversion. Hence, the operators L^2, L_z, S^2, S_z, $\mathcal{J}^2$, $\mathcal{J}_z$ and parity are constants of the motion which can be determined simultaneously. The interaction (13) (with a much less sophisticated spatial dependence) is one often assumed in nuclear physics. With it, the ground states of the deuteron, of ^{3}He and of ^{4}He are ${}^3S_1$, ${}^2S_{\frac{1}{2}}$ and ${}^1S_0$ states respectively.

One knows that the forces between protons and neutrons are not adequately represented by the central potential in the Hamiltonian H_1, equation (13). One must supplement it by a tensor interaction which contributes an additional term to the Hamiltonian which now becomes

$$H_2 = H_1 + \sum_{i<j}^{N} W_{ij} S_{ij} \tag{14}$$

where the W_{ij} are arbitrary operators which are invariant under separate rotations in orbital and spin space, and S_{ij} is the tensor operator

$$S_{ij} = \frac{3(\boldsymbol{\sigma}_i\mathbf{r}_{ij})(\boldsymbol{\sigma}_j\mathbf{r}_{ij})}{\mathbf{r}_{ij}^2} - (\boldsymbol{\sigma}_i\boldsymbol{\sigma}_j) \tag{15}$$

with $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. Under separate rotations in orbital and spin space, the operator S_{ij}, and hence the Hamiltonian H_2, is no longer invariant. For the scalar products $(\boldsymbol{\sigma}_i\mathbf{r}_{ij})$ are, of course, only invariant if we rotate axes for both vectors. Hence, with the Hamiltonian H_2, the operators L^2, L_z, S^2 and S_z are no longer constants of the motion.

The case of a two-particle system is an exception to the last result. In this case S^2 is a good quantum number even when using tensor forces. For the total Hamiltonian H_2 is now symmetric in the spin of the two particles, *i.e.* $[H_2, P_{12}^\sigma] = 0$. Hence the wave function must be either symmetric or antisymmetric in the spin coordinates, *i.e.* the spin wave functions must correspond to either a triplet or singlet state for which $S = 1$ and 0 respectively. Though S is a good quantum number in this case, S_z is not.

We obtain another example of a Hamiltonian which is only invariant under total rotations, from the field of atomic physics, where the central force Hamiltonian H_0, equation (41.2), is supplemented by a term representing the spin-orbit interaction. This term is usually taken proportional to

$$\mathbf{LS} \quad \text{or} \quad \sum_{i=1}^{N} \mathbf{l}_i\mathbf{s}_i \tag{16}$$

and is clearly not invariant under separate orbital and spin space rotations.

43. The representation of groups

Suppose Γ is a Hermitian operator, corresponding to a physical observable, which is invariant under a group of transformations. It follows from section 42, that Γ commutes with the group of operators $\mathfrak{G} = \{S, T, \ldots\}$,

$$[T, \Gamma] = 0, \quad \textit{etc.} \tag{1}$$

Suppose γ is an N-fold degenerate eigenvalue of Γ, *i.e.* there exists an N-dimensional manifold $\mathfrak{M}$ in Hilbert space such that all vectors of $\mathfrak{M}$ are eigenvectors of Γ belonging to the eigenvalue γ. If φ is any vector of $\mathfrak{M}$ and T an operator of the group $\mathfrak{G}$, it follows from equation (1) that

$$\Gamma(T\varphi) = T\Gamma\varphi = T\gamma\varphi = \gamma(T\varphi) \tag{2}$$

i.e. $(T\varphi)$ is also an eigenvector of Γ belonging to the *same* eigenvalue γ, *i.e.* it also lies in $\mathfrak{M}$. Hence $\mathfrak{M}$ is an *invariant* manifold of the group $\mathfrak{G}$: *operators of $\mathfrak{G}$ transform vectors of $\mathfrak{M}$ into vectors of $\mathfrak{M}$.*

If we choose any orthonormal set of vectors $\xi_1, \xi_2, \ldots \xi_N$ to span $\mathfrak{M}$, we can form the matrices

$$T_{ij} = (\xi_i, T\xi_j) \tag{3}$$

corresponding to the operators of the group $\mathfrak{G}$, and we call this set of $N \times N$ matrices, an N-dimensional *matrix representation* Δ of the group $\mathfrak{G}$. It has the property that to every operator T of the group corresponds a matrix T_{ik}, and to the product operator ST corresponds the product matrix $\Sigma S_{i\lambda} T_{\lambda k}$. Corresponding to a given group $\mathfrak{G}$, there will of course in general be several different matrix representations.

We saw in section 22, that we can choose many different orthonormal bases in $\mathfrak{M}$, and two such bases are related by a unitary transformation U such that $\tilde{T}_{ik} = (U^* T U)_{ik}$ [*cf* equation (22.13)]. The representations referred to two such bases will of course appear different, but they are not really significantly different: the effect of an operator T on a vector φ of $\mathfrak{M}$ is independent of the choice of basis. We call two such representations, differing only in the choice of bases, *equivalent* matrix representations.

If a vector φ belongs to the invariant manifold $\mathfrak{M}$, in which a representation Δ of the group $\mathfrak{G}$ occurs, we shall say that φ belongs to or transforms according to the representation Δ. The stress is here on the manifold $\mathfrak{M}$, independently of the choice of basis, that is, we treat equivalent representations essentially as a single representation.

Suppose the invariant N-dimensional manifold $\mathfrak{M}$ can be subdivided into two manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$, of dimensionality m and n respectively,

$$\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2 \quad (N = m + n) \tag{4}$$

and each of these two manifolds, by itself, is invariant under the group $\mathfrak{G}$, *i.e.* operators of $\mathfrak{G}$ transform vectors of $\mathfrak{M}_1$ into vectors of $\mathfrak{M}_1$, and similarly for $\mathfrak{M}_2$. [For a group of unitary operators, one can show that if $\mathfrak{M}_1$ is an invariant subspace of $\mathfrak{M}$, so is $\mathfrak{M}_2 = \mathfrak{M} - \mathfrak{M}_1$, *cf* Exercise IX.5.] If such a decomposition of $\mathfrak{M}$ is possible we call $\mathfrak{M}$ a *reducible* manifold, and the corresponding representation Δ *reducible*, otherwise we call the manifold $\mathfrak{M}$ and the representation Δ *irreducible*. We can now try and reduce, *i.e.* decompose into invariant subspaces, both the manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$. We can continue this process of reduction until $\mathfrak{M}$ is completely reduced, *i.e.* in the decomposition

$$\mathfrak{M} = \mathfrak{M}_1 + \mathfrak{M}_2 + \dots + \mathfrak{M}_r \tag{5}$$

all the sub-manifolds $\mathfrak{M}_1, \dots \mathfrak{M}_r$ are irreducible. Suppose we take a vector φ of an invariant manifold $\mathfrak{M}$ and operate on it with all the operators $\{S, T, \dots\}$ of the group. In this way we generate a manifold $\mathfrak{M}'$, spanned by the set of vectors $\{S\varphi, T\varphi, \dots\}$. If $\mathfrak{M}$ is irreducible, then $\mathfrak{M}' \equiv \mathfrak{M}$, for *any* vector φ of $\mathfrak{M}$ with which we start. If $\mathfrak{M}$ is reducible, then for *some* vectors φ of $\mathfrak{M}$, $\mathfrak{M}'$ will be a subspace of $\mathfrak{M}$. We can put this differently and say that only if $\mathfrak{M}$ is irreducible, will the group $\mathfrak{G}$ generate the *whole* manifold $\mathfrak{M}$ from *any* vector φ.

Suppose we span the reducible manifold $\mathfrak{M}$, equation (4), by the orthonormal vectors $\xi_1, \dots \xi_N$ which we choose so that the vectors

$\xi_1, \ldots \xi_m$ and $\xi_{m+1}, \ldots \xi_N$ span the manifolds $\mathfrak{M}_1$ and $\mathfrak{M}_2$ by themselves. Referred to this basis, any operator T of $\mathfrak{G}$ will have the matrix

$$\left[\begin{array}{c:c} \begin{matrix} T_{11} & \cdots\cdots & T_{1m} \\ T_{21} & \cdots\cdots & \\ \cdots & \cdots & \cdots \\ T_{m1} & \cdots\cdots & T_{mm} \end{matrix} & 0 \\ \hdashline 0 & \begin{matrix} T_{m+1,\,m+1} & \cdots & T_{m+1,\,N} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ T_{N,\,m+1} & \cdots\cdots & T_{NN} \end{matrix} \end{array}\right] \qquad (6)$$

since for $1 \leq i \leq m$,

$$T\xi_i = \sum_{k=1}^{N} \xi_k T_{ki} = \sum_{k=1}^{m} \xi_k T_{ki}, \quad \textit{i.e. } T\xi_i \text{ lies in } \mathfrak{M}_1,$$

and for $m+1 \leq i \leq N$,

$$T\xi_i = \sum_{k=1}^{N} \xi_k T_{ki} = \sum_{k=m+1}^{N} \xi_k T_{ki}, \quad \textit{i.e. } T\xi_i \text{ lies in } \mathfrak{M}_2.$$

Hence a reducible representation is of the form (6), with a suitable choice of basis in $\mathfrak{M}$. However, a unitary transformation to a basis in which the basis vectors are linear combinations of the $\xi_1, \ldots \xi_m$ with the $\xi_{m+1} \ldots \xi_N$ will spoil this appearance.

We shall illustrate our results for some of the groups introduced in section 39.

Let us consider the inversion group $\mathfrak{G}_I$ consisting of only two transformations, the identity $e:\mathbf{r} \to \mathbf{r}$ and the inversion $i:\mathbf{r} \to -\mathbf{r}$. The corresponding operators E, I applied to a wave function $\varphi(\mathbf{r})$ usually generate a two-dimensional manifold $\mathfrak{M}$ spanned by

$$E\varphi(\mathbf{r}) = \varphi(\mathbf{r}), \quad \text{and} \quad I\varphi(\mathbf{r}) = \varphi(-\mathbf{r}). \qquad (7)$$

This manifold is, however, always reducible into two one-dimensional (and hence necessarily irreducible) subspaces. For we can always write $\varphi(\mathbf{r})$ as the sum of a function of even and a function of odd parity (or parity ± 1 respectively),

$$\varphi(\mathbf{r}) = \tfrac{1}{2}\{\varphi(\mathbf{r}) + \varphi(-\mathbf{r})\} + \tfrac{1}{2}\{\varphi(\mathbf{r}) - \varphi(-\mathbf{r})\}, \qquad (8)$$

and a function $f(\mathbf{r})$ of definite parity $(-1)^n$ defines a one-dimensional manifold under reflection:

$$Ef(\mathbf{r}) = f(\mathbf{r}), \quad If(\mathbf{r}) = f(-\mathbf{r}) = (-1)^n f(\mathbf{r}). \qquad (9)$$

So we obtain two irreducible one-dimensional representations of the inversion group, the symmetric and the antisymmetric representations, for which $I = \pm 1$ respectively; we shall call them the $(+1)$ and (-1) representations.

Exactly the same two irreducible representations are obtained for the permutation group $\mathfrak{S}_2$ of two objects, as the reader will easily verify, the permutation operator P_{12} taking the place of the inversion operator I.

We next consider the rotation group which, as we know, is closely related to the angular momentum properties of a system, and one easily sees that the $(2n+1)$-dimensional manifold $\mathfrak{M}_n$, spanned by the eigenfunctions $\Psi(q, n, -n), \ldots \Psi(q, n, n)$ (*cf* section 25) of the angular momentum operators is invariant under the rotation group. For we saw, in section 25, that this manifold $\mathfrak{M}_n$ is invariant under the infinitesimal rotations (in fact, we generated it by repeated application of the infinitesimal rotation operators) and any finite rotation can be built up from these. One can also prove that the manifolds $\mathfrak{M}_n$, $n=0, \frac{1}{2}, 1, \ldots$, (and the corresponding representations, usually denoted by $\mathfrak{D}_n$), are irreducible and that there are no other irreducible representations of the rotation group. For n an integer, the wave functions $\Psi(q, n, m)$ are essentially the spherical harmonics $Y_n^m(\theta, \varphi)$ [*cf* equation (25.26)]. The representation $\mathfrak{D}_0$ is the identical one-dimensional representation corresponding to a wave function Y_0^0 which is independent of the angles (θ, φ) and hence invariant under rotations. The manifold $\mathfrak{M}_1$ of the three-dimensional representation $\mathfrak{D}_1$ is spanned by $Y_1^m(\theta, \varphi)$, $m=-1, 0, 1$, *i.e.* by x, y and z [*cf* equation (13.26)].

The addition of angular momenta in section 27 affords a useful example of the concept of reducibility. Equation (27.7),

$$\mathfrak{M}_{n_1}(1) \times \mathfrak{M}_{n_2}(2) = \sum_{n=n_1-n_2}^{n_1+n_2} \mathfrak{M}_n, \quad (n_1 \geqslant n_2) \tag{10}$$

in particular, corresponds to the reduction of the product manifold (which is itself invariant under $\mathfrak{R}_3$ and furnishes the product representation $\mathfrak{D}_{n_1} \times \mathfrak{D}_{n_2}$) into its irreducible sub-manifolds. One also often writes equation (10) for the representations themselves

$$\mathfrak{D}_{n_1} \times \mathfrak{D}_{n_2} = \sum_{n=n_1-n_2}^{n_1+n_2} \mathfrak{D}_n. \tag{11}$$

An analogous situation holds generally. Corresponding to two representations Δ_1 and Δ_2 of a group, one can form a product representation $\Delta_1 \times \Delta_2$ which is generally reducible. For example, the orbital wave function of two particles with orbital angular momentum n_1 and n_2 lies in the manifold (10). Under inversions, such a wave function transforms according to the product representation $(-1)^{n_1} \times (-1)^{n_2}$, *i.e.* according to the irreducible representation $(-1)^{n_1+n_2}$.

Lastly, we obtain the irreducible representations of the axial rotation group. This is an Abelian group [*cf* equation (39.3)] and as such its irreducible representations are necessarily one-dimensional (*cf* Exercise IX.6) *i.e.* to each rotation of axes through an angle α (we denoted this transformation by t_α in section 39) corresponds a 1×1 matrix, *i.e.* a single (complex) number $\varphi(\alpha)$. From equation (39.3) and the properties of a representation it then follows that

$$\varphi(\alpha)\varphi(\beta) = \varphi(\alpha+\beta). \tag{12}$$

The only continuous solutions of this functional equation are

$$\varphi(\alpha) = \exp(c\alpha), \quad (c = \text{a constant}) \tag{13}$$

and if this representation is to be single-valued, as we require in practice, we must have $\varphi(0) = \varphi(2\pi)$, *i.e.*

$$c = im, \quad (m = 0, \pm 1, \pm 2, \ldots). \tag{14}$$

Hence, the axial rotation group possesses the following irreducible representations (and one can show that they are the only ones)

$$\varphi_m(\alpha) = \exp(im\alpha), \quad (m = 0, \pm 1, \pm 2, \ldots). \tag{15}$$

So far we have only considered the representations of groups in terms of wave functions. But we have seen that a group of coordinate transformations $\{s, t, \ldots\}$ induces a group of transformations $\{S, T, \ldots\}$ in operators as well. Hence we can also talk of an operator A transforming according to a particular representation Δ_A of a group $\mathfrak{G}$, and of the wave function $A\varphi$ transforming according to the product representation $\Delta_A \times \Delta$, Δ being the representation to which φ belongs. For example, the interaction between two identical particles is necessarily unchanged if we interchange the particles, *i.e.* this interaction transforms according to the $(+1)$ or symmetric representation of the permutation group. The Coulomb repulsion e^2/r_{12} is an example of this. The 'electric dipole operator' (which we shall meet again in connection with selection rules)

$$D_z = \Sigma e_i z_i \tag{16}$$

of a system of electric charges e_i, with position coordinates $\mathbf{r}_i$, transforms according to the antisymmetric or (-1) representation of the reflection group $\mathfrak{G}_I$. Under rotations in spin-space $\mathfrak{S}$, the electric dipole operator is invariant, *i.e.* it transforms according to $\mathfrak{D}_0$; under rotations in orbital space $\mathfrak{H}$, it transforms according to $\mathfrak{D}_1$, since it transforms in the same way as the wave function $Y_1^0(\theta, \varphi) \propto z$. Applied to a wave function ψ of orbital angular momentum L, $D_z\psi$ transforms according to the product representation

$$\mathfrak{D}_1 \times \mathfrak{D}_L = \begin{cases} \mathfrak{D}_1, & \text{if } L = 0, \\ \mathfrak{D}_{L+1} + \mathfrak{D}_L + \mathfrak{D}_{L-1}, & \text{if } L > 1, \end{cases} \tag{17}$$

just like a wave function $Y_1^m Y_L^M$.

44. Perturbations

In Chapter VII we considered time-independent perturbation theory and, in particular, its application to the calculation of the shifting and splitting of energy levels by perturbing potentials. The group-theoretical methods, which we have so far developed, allow several quite fundamental deductions to be made in this field.

Suppose the Hamiltonian H^0 is invariant under a group $\mathfrak{G}$. If this Hamiltonian possesses an eigenvalue E^0 which is n_0-fold degenerate, the corresponding eigenfunctions span an n_0-dimensional manifold $\mathfrak{M}^0$ in which a representation of the group $\mathfrak{G}$ takes place, *i.e.* $\mathfrak{M}^0$ is invariant under the group $\mathfrak{G}$.

Let us consider the effect of a perturbation V which is also invariant under $\mathfrak{G}$. We vary the Hamiltonian and consider the Hamiltonian

$$H = H^0 + \varepsilon V, \tag{1}$$

obtained from H^0 by continuously increasing the parameter ε from $\varepsilon = 0$. As a result of varying the Hamiltonian, the degenerate energy level E^0 will be displaced and, in general, split into several levels, and the manifold $\mathfrak{M}^0$ will be transformed into a new manifold

$$\mathfrak{M} = \mathfrak{M}(\varepsilon) \tag{2}$$

comprising the corresponding perturbed eigenfunctions. If n is the dimensionality of $\mathfrak{M}$, then

$$n = n_0, \tag{3}$$

for we cannot produce a discontinuous change in n by increasing ε continuously from $\varepsilon = 0$. Since we can in this way make ε as large as we please, it follows that equation (3) is true generally for an arbitrarily large perturbation.

In general a perturbation splits a degenerate level. However, by means of the group-theoretical ideas which we have developed, we can obtain some restrictions on the number of levels into which a degenerate level can be split by a perturbation with the same symmetry properties as the unperturbed Hamiltonian. We shall firstly prove the important result that if $\mathfrak{M}^0$ is irreducible, so is $\mathfrak{M}$. For suppose that for some small positive value of ε, $\mathfrak{M}$ is reducible. The matrices of the operators of $\mathfrak{G}$, referred to a suitable basis, are then of the form

$$\left[\begin{array}{c|c} A(\varepsilon) & 0 \\ \hline 0 & B(\varepsilon) \end{array}\right] \tag{4}$$

where $A(\varepsilon)$ and $B(\varepsilon)$ are submatrices whose elements depend of course on ε. But if we proceed to the limit $\varepsilon \rightarrow 0$, a matrix of the form (4) always retains this form, *i.e.* in the limit it is of the form

$$\left[\begin{array}{c|c} A(0) & 0 \\ \hline 0 & B(0) \end{array}\right] \tag{5}$$

i.e. $\mathfrak{M}^0$ must also be reducible, contrary to hypothesis. As before, we can make the perturbation arbitrarily large by increasing ε continuously. We then have

Theorem 1. A perturbation V with the same invariance properties as the unperturbed Hamiltonian H^0 can never split a degenerate level corresponding to an irreducible representation, however large the perturbation may be.

The perturbation will of course in general displace the energy level. We find the displaced energy level, to the first order in ε, from equation (29.13), to be

$$E = E^0 + \varepsilon(\psi, V\psi) \tag{6}$$

where ψ is *any* normed function lying in the manifold $\mathfrak{M}^0$.

We can verify that the matrix element $(\psi, V\psi)$ has the same value for any vector ψ of $\mathfrak{M}^0$ (and so again prove theorem 1) by considering two different normed functions ψ and φ of $\mathfrak{M}^0$. Since $\mathfrak{M}^0$ is irreducible, there exists a unitary operator T of the group $\mathfrak{G}$ such that $\varphi = T\psi$ and $[T, V] = 0$. Hence

$$(\psi, V\psi) = (T\psi, TV\psi) = (T\psi, VT\psi) = (\varphi, V\varphi). \tag{7}$$

The degeneracy associated with an irreducible manifold is called *essential.*

We must now consider the case where the manifold $\mathfrak{M}^0$ is reducible. In this case we decompose the manifold $\mathfrak{M}^0$ into its irreducible manifolds

$$\mathfrak{M}^0 = \mathfrak{M}_1 + \mathfrak{M}_2 + \ldots + \mathfrak{M}_s, \tag{8}$$

where $\mathfrak{M}_i$ $(i = 1, \ldots s)$ is an irreducible space of dimensionality n_i $(n = n_1 + n_2 + \ldots + n_s)$. Theorem 1 then at once gives

Theorem 2. If the perturbation V has the same invariance properties as the unperturbed Hamiltonian H^0, and the manifold $\mathfrak{M}^0$ of the unperturbed level E^0 can be completely reduced, according to (8), then the perturbation V, however large, can at most split the level E^0 into s levels $E_1, E_2, \ldots E_s$ of degeneracy $n_1, n_2, \ldots n_s$ respectively.

We call this degeneracy, which can be split by a perturbation V with the same symmetry as H^0, *accidental.*

To find the energy level E_i to the first order in ε, we must decompose $\mathfrak{M}^0$ into its irreducible subspaces, equation (8). We obtain E_i, as before, by taking *any* normed wave function ψ_i belonging to the manifold $\mathfrak{M}_i$; then,

$$E_i = E^0 + \varepsilon(\psi_i, V\psi_i), \quad (i = 1, \ldots s). \tag{9}$$

Equation (9) gives the energy level E_i correctly only if the manifold $\mathfrak{M}_i$ is inequivalent to all the other manifolds occurring in the decomposition (8), as is shown below*.

[By a slight generalization of section 43, we call two invariant manifolds $\mathfrak{M}_i$ and $\mathfrak{M}_j$ equivalent if they have the same dimensionality $n_i = n_j$, and we can span them with orthonormal vectors $\xi_1, \xi_2, \dots \xi_{n_i}$ and $\eta_1, \eta_2, \dots \eta_{n_i}$ such that the matrices of the operators $\{S, T, U, \dots\}$ referred to these bases are the same

$$(\xi_l, T\xi_k) = (\eta_l, T\eta_k), \quad (k, l = 1, \dots n_i).$$

If this is not possible, we call the manifolds, and the corresponding representations, inequivalent.

Suppose the invariant manifolds $\mathfrak{M}_i$ and $\mathfrak{M}_j$ are equivalent. In this case the further decomposition of the manifold $\mathfrak{M}_i + \mathfrak{M}_j$ is not completely determined. In fact, we can reduce it into two irreducible mutually orthogonal manifolds in infinitely many ways. Any two sets of functions

$$\left.\begin{aligned} f_{k1} &= a\xi_k + b\eta_k \\ f_{k2} &= c\xi_k + d\eta_k \end{aligned}\right\} \qquad (k = 1, 2, \dots n_i) \tag{10}$$

which are mutually orthogonal, $(f_{k1}, f_{k2}) = 0$, define two such manifolds. We must then use the methods of section 30 to determine the correct linear combinations of f_{k1} and f_{k2} such that the matrix (f_{ki}, Vf_{kj}) becomes diagonal. But we need only do this for any one value of k; for we see from (10), that the same linear combinations then hold for all k. By choosing the linear combinations (10), we have thus only to diagonalize a (2×2) matrix, instead of a $(2n_i \times 2n_i)$-matrix. If, more generally, there are ν equivalent manifolds $\mathfrak{M}_i, \mathfrak{M}_{i+1} \dots \mathfrak{M}_{i+\nu-1}$, we have to diagonalize a $\nu \times \nu$ matrix.]

Let us illustrate the above results by some simple examples.

(*1*) Consider an atom with Z electrons. Neglecting the mutual interaction of the electrons, the atom will be described by a product wave function [*cf* equation (25.26)]

$$\psi(q_1, l_1, m_1) \dots \psi(q_Z, l_Z, m_Z)$$

and will hence have the parity $(-1)^{l_1+l_2+\dots\dots+l_Z}$. Introducing an arbitrarily large interaction V, which is invariant under inversion (*e.g.* the Coulomb repulsion $\sum_{i<j} e^2/r_{ij}$), does not change this parity, even though the atomic wave function is no longer just a product of one-electron wave functions. For introducing the perturbation continuously cannot produce a discontinuous change by a factor (-1) in the parity.

---

* The following discussion, in parenthesis, will not be needed, and may be omitted if the reader so desires.

(*2*) For the problem of two identical particles, which we have repeatedly treated, any perturbation V must of course also be invariant under the permutation of the particles. Hence it cannot split the exchange degeneracy. The correct linear combinations for calculating the perturbation energy are the symmetric and antisymmetric wave functions which each 'span' a one-dimensional irreducible manifold. Hence, the first-order corrections to the energy levels are given by

$$\tfrac{1}{2}([u_1(1)u_2(2) \pm u_1(2)u_2(1)],\quad V[u_1(1)u_2(2) \pm u_1(2)u_2(1)]) = V_{11} \pm V_{12},$$

agreeing with our previous result, equation (31.18), but obtained with much less trouble.

(*3*) Consider two atomic electrons with orbital angular momenta l_1 and $l_2(\leq l_1)$. Neglecting their interaction, the energy level E^0 and its manifold $\mathfrak{M}$ of eigenvalues are $2(2l_1+1)(2l_2+1)$-fold degenerate. We want to calculate the splitting of this level due to the Coulomb repulsion e^2/r_{12} between the electrons. This is invariant under both spatial rotations and permutations. To begin with, we reduce the manifold $\mathfrak{M}$ with respect to the rotation group $\mathfrak{R}_3$, and obtain $(2l_2+1)$ $2(2L+1)$-dimensional invariant manifolds $\mathfrak{M}_L$ of $\mathfrak{R}_3$, with $L = l_1 - l_2, \ldots, l_1 + l_2$. The manifold $\mathfrak{M}_L$ corresponds to states of resultant orbital angular momentum L, and it is, of course, not irreducible. It is spanned by the wave functions

$$\varphi_{L\pm}^{M} = \sum_{m_1+m_2=M} C_{l_1m_1l_2m_2}^{LM} \frac{1}{\sqrt{2}} \{\Psi_1(l_1, m_1)\Psi_2(l_2, m_2) \pm \Psi_2(l_1, m_1)\Psi_1(l_2, m_2)\},$$
$$(M = -L, \ldots L), \tag{11}$$

where the C's are Wigner coefficients [*cf* equation (27.13)]. Since there exists only one irreducible representation of $\mathfrak{R}_3$, corresponding to wave functions of angular momentum L, namely $\mathfrak{D}_L$, it follows that $\mathfrak{M}_L$ must consist of two such equivalent manifolds. However, if we now consider the reduction of the manifold $\mathfrak{M}_L$ under the symmetric group $\mathfrak{S}_2$, we at once obtain two $(2L+1)$-dimensional manifolds $\mathfrak{M}_{L\pm}$. Each of these is irreducible under $\mathfrak{R}_3$ and $\mathfrak{S}_2$, transforming according to the $\mathfrak{D}_L$ and (± 1) representations of these two groups respectively, and being spanned by the wave functions $\varphi_{L\pm}^{M}$ $(M = -L, \ldots L)$, respectively. The Coulomb interaction then can at most split the $2(2l_1+1)(2l_2+1)$-fold level E^0 into $2(2l_2+1)$ levels (and, in general, will do so), given by

$$E_{L\pm} = E^0 + (\varphi_{L\pm}^{M}, V\varphi_{L\pm}^{M}), \quad (L = l_1 - l_2, \ldots l_1 + l_2). \tag{12}$$

The level $E_{L\pm}$ is $(2L+1)$-fold degenerate, and we can take any one of the values of $M(= -L, \ldots L)$ in calculating it.

We can see the M-independence of the matrix element $(\varphi_{L\pm}^{M}, V\varphi_{L\pm}^{M})$ in yet another way. The interaction V is invariant under rotations and so independent of the azimuthal angles φ_1 and φ_2 of the electrons. These angles enter the wave function φ_{LS}^{M} only through exponentials like $\exp i[(m_1\varphi_1)+(M-m_1)\varphi_2]$, *i.e.* the matrix element in (12) through terms like $\exp i[(m_1-m_1')(\varphi_1-\varphi_2)]$. These are independent of M, and furthermore vanish unless $m_1=m_1'$, *i.e.* the 'cross-terms' do not contribute to the matrix element.

45. Selection rules

In Chapter VIII we considered collision problems and obtained as the general result certain transition probabilities w which were proportional to the square of the modulus of matrix elements of the form

$$(\varphi', A\varphi). \tag{1}$$

Here φ' represents some 'final' state in which we are particularly interested, such as one in which a particle is scattered into a given direction, A is essentially the interaction which causes the scattering, and φ is the actual wave function of the system. In general φ is not known and we approximate it somehow, for example, by a wave function of the unperturbed Hamiltonian, *i.e.* disregarding the interaction A.

To calculate a transition probability one must of course evaluate the matrix element (1). But in many cases it is possible to deduce by general symmetry arguments that a particular interaction A cannot cause transitions to a particular final state φ': the matrix element (1) must vanish. Such rules, forbidding certain transitions, are known as *selection rules*. Though one can of course obtain such selection rules without introducing any group theory explicitly, this procedure does bring out the essential ideas common to all selection rules.

To obtain selection rules for a particular process, we consider the symmetry groups of the system. For angular momentum selection rules, this is the rotation group, for parity selection rules it is the reflection group.

Consider a transformation group $\mathfrak{G}$. In general it will be possible to split the function space $\mathfrak{F}$, in which our system is described quantum mechanically, into mutually orthogonal subspaces $\mathfrak{M}_1, \mathfrak{M}_2, \ldots \mathfrak{M}_i \ldots$ which, together, span the whole function space,

$$\mathfrak{F}=\mathfrak{M}_1+\mathfrak{M}_2+\ldots+\mathfrak{M}_i+\ldots \tag{2}$$

and these subspaces $\mathfrak{M}_1, \mathfrak{M}_2, \ldots \mathfrak{M}_i, \ldots$ have the property of being invariant and irreducible under the group $\mathfrak{G}$, *i.e.* certain irreducible representations $\Delta_1, \Delta_2, \ldots \Delta_i, \ldots$ of the group $\mathfrak{G}$ will occur in them. (That these manifolds are mutually orthogonal is related to the fact that they belong to different eigenvalues of some operator, such as the angular momentum L^2.)

We can expand any wave function ψ into components belonging to the different manifolds $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ occurring in (2), *i.e.* into components transforming according to the irreducible representations $\Delta_1, \Delta_2, \ldots,$

$$\psi = \xi_1 + \xi_2 + \ldots + \xi_i + \ldots, \tag{3}$$

where ξ_i lies in $\mathfrak{M}_i$ and is uniquely determined.

Under the group $\mathfrak{G}$, the wave functions φ, φ' and the operator A will transform according to some representations Δ, Δ' and Δ_A of this group, respectively. In general, these representations will of course be reducible. The wave function

$$\chi = A\varphi \tag{4}$$

will transform according to the product representation

$$\Delta_A \times \Delta \tag{5}$$

which, of course, will also in general be reducible. We now expand the wave functions χ and φ' into components belonging to different irreducible manifolds $\mathfrak{M}_1$, $\mathfrak{M}_2, \ldots,$ as indicated in equation (3). It then follows that the matrix element

$$(\varphi', \chi) = (\varphi', A\varphi) \tag{6}$$

must vanish, if the expansions of χ and φ' contain no common manifold $\mathfrak{M}_i$, for in this case the two wave functions are necessarily orthogonal. Only if these expansions contain at least one common manifold $\mathfrak{M}_i$ can the matrix element be different from zero. We can express this differently by saying that the matrix element (6) vanishes, if the complete reductions of the representations $\Delta_A \times \Delta$ and Δ' contain no irreducible representation Δ_i in common. We can interpret this result as follows. The matrix element (6), corresponding to a physically observable transition probability, must be independent of the coordinate system chosen, that is, it must be invariant under the group of transformations $\mathfrak{G}$. Hence, only those parts of the wave functions χ and φ' can give a non-vanishing contribution to the matrix element, which transform in the same manner (*i.e.* according to the same irreducible representations). The matrix elements of parts of the wave functions which transform differently must vanish identically, as otherwise they would give contributions to the matrix element which change under transformations of the group. The application of these methods is greatly simplified in many cases by the fact that the representations Δ_A, Δ and Δ', according to which A, φ and φ' transform, are irreducible. We need, then, only investigate whether the product representation $\Delta_A \times \Delta$ contains the representation Δ'. It is in general only in this case that enough transitions are forbidden for any useful selection rules to arise.

To deal with parity selection rules, we consider the inversion group

$\mathfrak{G}_I$. If φ, φ' and A transform according to the irreducible representations $(-1)^n$, $(-1)^{n'}$ and $(-1)^a$, then $A\varphi$ transforms according to the irreducible representation $(-1)^{a+n}$, and transitions can only occur if

$$(-1)^{a+n} = (-1)^{n'}$$

that is

$$a + n + n' = \text{even}. \qquad (7)$$

To obtain angular momentum selection rules we consider the rotation group. If, under rotations in spin space, φ, φ' and A transform according to the irreducible representations $\mathfrak{D}_S$, $\mathfrak{D}_{S'}$ and $\mathfrak{D}_{S_A}$, then $A\varphi$ transforms according to

$$\mathfrak{D}_{S_A} \times \mathfrak{D}_S = \mathfrak{D}_{|S_A - S|} + \ldots + \mathfrak{D}_{S_A + S} \qquad (8)$$

and the matrix element (6) vanishes unless S' is equal to one of the numbers $|S_A - S|, \ldots, S_A + S$. In particular, if A transforms according to $\mathfrak{D}_0$, *i.e.* is invariant under spin rotations, then transitions can only occur if $S = S'$. Since $(2S+1)$ and $(2S'+1)$ are the spin multiplicities of initial and final states, corresponding to resultant spins S and S', $S = S'$ means that the spin multiplicity has not been changed by the transition. Any operator A which only operates on the orbital part of the wave function has of course this property. In exactly the same way one obtains selection rules for orbital and total angular momenta, provided the initial and final states and the operator A transform according to irreducible representations of the orbital and total rotation groups.

The methods which we have given for obtaining selection rules are very useful and can be applied to a great variety of problems: radiative processes in atoms and nuclei, β-decay and nuclear reactions generally. The reader will find most of these discussed in the books by Sommerfeld[13] and Blatt and Weisskopf[20]. We shall only give a few illustrative examples.

(*I*) The line spectra of atoms are due to radiative transitions, in which an atom in an excited state φ, with energy E, 'drops' to a less excited state φ', with energy $E' < E$. At the same time a quantum of radiation of frequency

$$\nu = (E - E')/h$$

is emitted.

In the quantum theory of radiation it is shown that the transition probability of this process is proportional to the modulus squared of the matrix element

$$(\varphi', D_e\varphi) \qquad (9)$$

where D_e is the electric dipole operator of the atom

$$D_e = e\sum_{i=1}^{Z} \mathbf{r}_i \quad \leftarrow \text{2cap.} \qquad (10)$$

(Z = number of atomic electrons, $\mathbf{r}_i$ = position vector of ith electron).

The dipole operator D_z is the component of a vector. Hence it transforms according to the (-1) representation of the inversion group. It is independent of spin and hence transforms according to $\mathfrak{D}_0$ under spin rotations. Under orbital rotations it transforms according to $\mathfrak{D}_1$, being the component of a vector, and under total rotations according to $\mathfrak{D}_1 \times \mathfrak{D}_0 = \mathfrak{D}_1$. If φ and φ' have total angular momenta J and J' and parity $(-1)^{\pi}$ and $(-1)^{\pi'}$, then

$$\mathfrak{D}_1 \times \mathfrak{D}_J = \begin{cases} \mathfrak{D}_{J+1} + \ldots \mathfrak{D}_{|J-1|}, & \text{if } J \geq \tfrac{1}{2} \\ \mathfrak{D}_1, & \text{if } J = 0. \end{cases} \tag{11}$$

Hence equation (7) gives the parity selection rule

$$\pi' = \pi + 1 \tag{12}$$

i.e. an electric dipole transition changes the parity of the atomic state. This is known as *Laporte's rule*. The angular momentum selection rules, obtained from (11), are, if we define $\Delta J \equiv J' - J$, given by

$$\Delta J = 0, \pm 1 \quad (\textit{not } 0 \rightarrow 0) \tag{13}$$

with the transition $J = 0$ to $J' = 0$ excluded as follows from equation (11), as we indicated in brackets in equation (13).

If the total spin and total orbital angular momenta by themselves are good quantum numbers so that we can ascribe definite values S, L and S', L' to them in the two states, we obtain the further selection rules ($\Delta S \equiv S' - S$, $\Delta L \equiv L' - L$):

$$\Delta S = 0 \tag{14}$$

$$\Delta L = 0, \pm 1 \quad (\textit{not } 0 \rightarrow 0). \tag{15}$$

Equation (14) states that there are no spectral lines connecting states of different spin multiplicities (known as intercombination lines). Such lines are not observed for light atoms, such as helium, but do occur for heavy atoms. This is in agreement with the fact that the spin-orbit interaction is only important for heavy atoms, and it is of course a consequence of spin-orbit interaction that L and S are no longer good quantum numbers [*cf* equation (42.16)].

If one considers radiative transitions, due to an electric dipole interaction analogous to (10), in nuclei, one obtains the same selection rules as above. Selection rules (14) and (15) now only apply if we use central forces and neglect tensor forces [*cf* equation (42.15)].

(2) In addition to these radiative transitions due to the electric dipole operator (10), other types of radiative transitions can occur. These various types of transitions differ in the angular momentum and parity properties of the emitted radiation. However, these other kinds of radiative transitions are very weak compared with electric dipole radiation. Consequently, they only become important for transitions for which the matrix element of the electric dipole operator vanishes, on account of a selection rule. They are almost exclusively of interest in nuclear rather than atomic physics.

In nuclear physics, magnetic dipole transitions are of importance. For a nucleus consisting of $\mathcal{Z}$ protons and $\mathcal{N}$ neutrons, whose coordinates we label with the numbers $1, 2, 3, \ldots \mathcal{Z}$, and $\mathcal{Z}+1, \mathcal{Z}+2, \ldots \mathcal{Z}+\mathcal{N}=A$, respectively, the magnetic dipole transitions from a nuclear state φ to a state φ' are associated with the matrix element

$$\frac{e\hbar}{2Mc}\int d\tau\,\overline{\varphi'}\left\{\sum_{k=1}^{\mathcal{Z}} l_z(k)+\sum_{k=1}^{A}\mu_k\sigma_z(k)\right\}\varphi \tag{16}$$

where $l_z(k)$ and $\sigma_z(k)$ are the z-components of the orbital angular momentum and spin operators of the kth particle and μ_k is its intrinsic magnetic moment measured in units of $e\hbar/2Mc$ (mass of neutron $=$ mass of proton $=M$; $e=$ electric charge of proton; μ_k of course assumes only two values). One easily obtains the selection rules for magnetic dipole transitions (*cf* Exercise IX.7).

(*3*) It is of interest to indicate briefly how the interaction operators (10) and (16) come about. For simplicity, we consider the particular case of a single spinless particle of charge e. For a radiative transition in which the particle changes from a state φ to a state φ', emitting or absorbing a quantum of radiation of momentum $\hbar k$ (*i.e.* wavelength $2\pi/k$) in the z-direction with a polarization vector pointing in the x-direction, the matrix element, derived from quantum electrodynamics, is

$$\int \overline{\varphi'(\mathbf{r})}\, e p_x e^{\pm ikz}\varphi(\mathbf{r})\,d\mathbf{r} \tag{17}$$

where $p_x=-i\hbar\partial/\partial x$ is the operator of the x-component of momentum and the plus and minus signs respectively apply to processes in which a photon is absorbed or created.

To obtain parity and angular momentum selection rules relating the initial and final states φ and φ' in equation (17), we must analyse the exponential function into terms which themselves have definite parity and angular momentum properties. It can be shown that these are in fact just the properties of the photon involved in the particular process considered, so that the selection rules merely state the conservation of angular momentum and parity for the system as a whole, *i.e.* particle + radiation.

In most applications, either or both of the wave functions φ and φ' will refer to states in which the particle is restricted to a definite region; it may be an electron in an atom, or a proton or neutron in a nucleus. The wave functions φ and φ' are then only large in a region $r<R$, corresponding to the extent of the atom or nucleus. In practice the wavelength $\lambda=2\pi/k$ of the radiation is usually large compared with the size of the atom or nucleus, so that kR is very small. (For an atomic electron, and light in the visible part of the spectrum, we have $\lambda\simeq 6\times 10^{-5}$ cm, $R\simeq 10^{-8}$ cm, so that $kR\simeq 10^{-3}$. For x-rays for which

$\lambda \simeq 1$ to 10×10^{-9} cm our arguments do not apply. For nuclei, $R \simeq 2$ to 8×10^{-13} cm, so kR is small even for photons with several MeV. For a 3 MeV γ-ray and $R \simeq 6 \times 10^{-13}$ cm, corresponding to a nucleus like copper, $kR \simeq 0{\cdot}1$.) We can then expand the exponential in (17) into a power series and successive terms will decrease by a factor $1/kR$ approximately:

$$\mathrm{e}^{\pm ikz} = 1 \pm ikz + \dots \tag{18}$$

Substituting the first of these terms in (17) we obtain a term proportional to the matrix element (*cf* Exercise IX.8)

$$\int \overline{\varphi'(\mathbf{r})}(ex)\varphi(\mathbf{r})d\mathbf{r} \tag{19}$$

of the electric dipole interaction, equation (10). Hence, if our condition $kR \ll 1$ is fulfilled, radiative transitions are largely due to the electric dipole interaction.

If on account of selection rules, the matrix element (19) vanishes, one must consider the second term in (18) giving the matrix element

$$iek\int \overline{\varphi'(\mathbf{r})}(p_x z)\varphi(\mathbf{r})d\mathbf{r}. \tag{20}$$

The operator $(p_x z)$ in this matrix element has a definite parity but does not belong to one irreducible representation of the rotation group, so we must decompose it suitably.

If we consider two vectors $\mathbf{r}$ and $\mathbf{p}$, their 9-dimensional product space is spanned by all cross-products

$$xp_x, yp_x, zp_x, xp_y, \dots, zp_z,$$

and forms an invariant manifold transforming according to $\mathfrak{D}_1 \times \mathfrak{D}_1 = \mathfrak{D}_2 + \mathfrak{D}_1 + \mathfrak{D}_0$ under rotations. One easily finds the irreducible subspaces transforming according to $\mathfrak{D}_0$ and $\mathfrak{D}_1$. The invariant quantity one can form from two vectors is the scalar product

$$\mathbf{rp} = xp_x + yp_y + zp_z$$

and the vector (transforming according to $\mathfrak{D}_1$) is the vector product $\mathbf{r} \wedge \mathbf{p}$ with components $xp_y - yp_x$, *etc.* Comparing these expressions with the operator $p_x z$ occurring in (20) one sees that one must write it as, (z and p_x commute),

$$p_x z = \tfrac{1}{2}(zp_x - xp_z) + \tfrac{1}{2}(zp_x + xp_z). \tag{21}$$

The first of these expressions is simply the y-component l_y of the orbital angular momentum, and taking this term by itself leads to the magnetic dipole transitions (*cf* equation (16): we are now considering a spinless particle, hence the second term of equation (16) drops out). The second term in (21) transforms according to $\mathfrak{D}_2$ under rotations and leads to what are known as electric quadrupole transitions. In this way one can continue and obtain the selection rules for higher order transitions.

The method which we have here developed is also exceedingly useful in the theory of β-decay, *i.e.* the disintegration of a neutron (usually in a nucleus) into a proton and an electron. (In addition a very light particle of spin $\frac{1}{2}$, called a neutrino, must be produced in these disintegrations to ensure conservation of energy and angular momentum.) The matrix element for such a transition is of the form

$$\int \overline{\varphi'(\mathbf{r})}\, V e^{i\mathbf{K}\cdot\mathbf{r}} \varphi(\mathbf{r})\, d\mathbf{r} \tag{22}$$

where φ is the initial wave function of the neutron, φ' the final wave function of the proton and $\hbar\mathbf{K}$ the momentum transferred to the electron and neutrino. V represents an interaction potential which is not fully known. General considerations show that for slow neutrons and protons (so that they can be described non-relativistically) this interaction can have only two forms, either

$$V = g \tag{23a}$$

or

$$V = g'\boldsymbol{\sigma} \tag{23b}$$

where g and g' are constants and $\boldsymbol{\sigma}$ is the spin vector of the nucleon (*i.e.* the particle which starts life as a neutron and transforms to a proton). Again, the wave functions φ and φ' are only appreciable inside the nucleus, $r < R$, and the momentum transfer is usually such that $KR \ll 1$. Hence we can again expand the exponential in (22) and need consider the first non-vanishing term only. β-decays corresponding to the first term (1) are known as allowed transitions, successive terms as first, second, . . . forbidden transitions. The reader should have no difficulty in deriving the selection rules for these cases. (A shorter and less complete account of β-decay, than in Blatt and Weisskopf[20], containing the essentials however, will be found in Bethe and Morrison[19].)

EXERCISES IX

IX.1. Show that the permutation group of n objects, $\mathfrak{S}_n$, is Abelian, if $n = 2$, and is not Abelian if $n > 2$.

IX.2. If $\mathfrak{S}_n = \{s, t, u, \ldots\}$ is the group of permutations of n objects, prove that the corresponding group of operators $\{S, T, U, \ldots\}$ consists of unitary operators.

IX.3. Prove that each of equations (42.7) implies the corresponding equation (42.8) and vice versa.

IX.4. Given that the ground state of the deuteron is a $^3S_1$ state, under the assumption of central forces, how is this state modified if we assume some tensor force acting?

IX.5. If $\mathfrak{M}$ is an invariant manifold of a group of unitary operators $\mathfrak{G} = \{S, T, \ldots\}$ and $\mathfrak{M}_1$ is an invariant subspace of $\mathfrak{M}$, prove that the subspace of $\mathfrak{M}$ orthogonal to $\mathfrak{M}_1$, *i.e.* $\mathfrak{M}_2 = \mathfrak{M} - \mathfrak{M}_1$, is also invariant under the group $\mathfrak{G}$.

IX.6. Prove that the irreducible representations of an Abelian group of unitary operators are one-dimensional.

IX.7. Obtain the selection rules for magnetic dipole radiation, the matrix element for which is given by equation (45.16).

IX.8. Show that the matrix element of the electric dipole transition, equation (45.19), follows from equation (45.17) if one approximately puts $\exp(ikz) = 1$, *i.e.* takes the first term in equation (45.18).

IX.9. The matrix element for electric quadrupole transitions from a nucleus, of charge Ze and mass number A, is, by generalization from equations (45.20–21), found to be proportional to

$$\sum_{k=1}^{Z} \int \overline{\psi'} [z_k p_x(k) + x_k p_z(k)] \psi \, d\mathbf{r}_1 \ldots d\mathbf{r}_A$$

where $\mathbf{r}_k$ and $\mathbf{p}(k)$ are the position and momentum vectors of the kth proton and neutron for $1 \leq k \leq Z$ and $Z+1 \leq k \leq N$ respectively. Obtain the selection rules for electric quadrupole transitions.

X

RELATIVISTIC THEORY OF THE ELECTRON

46. The Dirac equation for a free particle

The quantum mechanics which has so far been developed in this book does not satisfy the requirements of the special theory of relativity. We see this, for example, by looking at the Schrödinger equation which was derived from a non-relativistic Hamiltonian. It is a first-order differential equation in time but second-order in the space coordinates: time and space coordinates do not enter it in a symmetric form—on an equal footing as we may say—as required by the special theory of relativity.

These remarks suggest two ways of obtaining relativistic wave equations. The first of these was already proposed by Schrödinger in his original papers on wave mechanics. Instead of the non-relativistic Hamiltonian for a free particle, equation (5.7), he takes the relativistic Hamiltonian

$$E^2 = m^2c^4 + c^2\mathbf{p}^2 \tag{1}$$

and now makes the operator replacements

$$E \rightarrow i\hbar\frac{\partial}{\partial t}, \quad \mathbf{p} \rightarrow -i\hbar \operatorname{grad} \tag{2}$$

[equations (6.1), (6.2)] in this equation. One obtains

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{m^2c^2}{\hbar^2}\right]\psi(\mathbf{x}, t) = 0. \tag{3}$$

This equation is known as the Klein–Gordon equation.

We can write it in a way which brings out its relativistic invariance* by using a four-vector notation. We shall define the four-vector x_μ to have components $(x_1, x_2, x_3, x_4 = ix_0 = ict)$. We shall also write this $(\mathbf{x}, x_4)$ and generally will use Greek subscripts, $\lambda, \mu, \nu, \ldots$ running from 1 to 4, to denote four-vectors, whereas Latin subscripts, $i, j, k, l, m, \ldots$ running from 1 to 3, will denote spatial coordinates. The scalar product of two four-vectors x_μ and y_μ is defined by

$$\sum_{\mu=1}^{4} x_\mu y_\mu. \tag{4a}$$

---

* By relativistic invariance we always mean invariance under Lorentz transformations.

We shall use the convention usual in relativity theory of writing this simply as

$$x_\mu y_\mu \tag{4b}$$

i.e. summation over *repeated* Greek suffixes in a product is implied unless stated otherwise.

The constant $(\hbar/mc)$ in equation (3), characteristic of the mass of the particle, has the dimensions of a length and is called the Compton wavelength of the particle. For simplicity we shall denote it by $1/M$. The Klein–Gordon equation then takes the simple form

$$(\Box^2 - M^2)\psi(x_\nu) = 0, \tag{5}$$

where

$$\Box^2 \equiv \frac{\partial^2}{\partial x_\mu \partial x_\mu} = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}. \tag{6}$$

Equation (5) is relativistically invariant, provided the wave function $\psi(x_\nu)$ is invariant under Lorentz transformations, *i.e.* provided it is a *scalar*, borrowing this term from three-dimensional vector analysis.

The use of the Klein–Gordon equation as the relativistic theory for electrons meets several difficulties. We shall not discuss these. It suffices for our purposes to see that equation (5) cannot be the correct relativistic equation for electrons. For equation (5) is a wave equation for a one-component wave function*, whereas we have seen that even non-relativistically a particle with spin $\frac{1}{2}$ requires a two-component wave function, corresponding to the two degrees of freedom of the spin. Indeed one can show that in the non-relativistic limit of small kinetic energies, the Klein–Gordon equation reduces to the Schrödinger equation without spin and not to the non-relativistic Pauli theory, developed in section 26. There we introduced the spin and the magnetic moment in an *ad hoc* manner. Going back to section 26, we would like to stress that, by analogy with equation (26.2), one would expect a particle of intrinsic spin $\mathbf{s}$ to possess an intrinsic magnetic moment (we take $e>0$ so that the electron has the charge $-e$)

$$\frac{-e}{2mc}\mathbf{s}$$

and not

$$\frac{-e}{mc}\mathbf{s}$$

as postulated in equation (26.4) and as required to agree with a vast amount of spectroscopic and other experimental data. It is one of the many virtues of the Dirac equation of the electron that it automatically gives the correct spin and magnetic moment.

---

* One could of course consider ψ as consisting of two or more scalar wave functions without violating relativistic invariance. These would then be completely independent of each other. Under three-dimensional rotations, for example, they would not transform correctly to represent a particle with spin (*cf.* Exercise VI.6 and section 42) and thus cannot represent an electron.

In going from the Schrödinger to the Klein–Gordon equation we obtained symmetry between space and time coordinates by having second-order derivatives throughout. Dirac had the ingenious idea of working with a relativistic equation linear in space and time coordinates. We shall write this equation, which is known as the Dirac equation,

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + M\right)\psi(x_\nu) = 0 \tag{7}$$

where the γ_μ are so far undetermined. [In equation (7), remember that summation over μ from 1 to 4 is understood.] The Dirac equation looks very different from the Klein–Gordon equation. We do of course still want the classical Hamiltonian, equation (1), to hold and for many purposes of interpretation one wants an equation resembling the Klein–Gordon equation. The characteristics of the Klein–Gordon equation are that it is of the second order in all derivatives and does not contain any first-order derivatives. We obtain such an equation from the Dirac equation by operating on it with the operator

$$\left(\gamma_\lambda \frac{\partial}{\partial x_\lambda} - M\right)$$

which gives

$$\left(\gamma_\lambda \gamma_\mu \frac{\partial^2}{\partial x_\lambda \partial x_\mu} - M^2\right)\psi(x_\nu) = 0. \tag{8}$$

This equation becomes identical with the Klein–Gordon equation if the γ_λ satisfy

$$\gamma_\lambda \gamma_\mu + \gamma_\mu \gamma_\lambda = 2\delta_{\lambda\mu}, \quad (\lambda, \mu = 1, \dots 4). \tag{9}$$

If the γs are ordinary numbers, these relations cannot make sense. Ordinary numbers do not anticommute. However, if one interprets the γs as matrices, it may be possible to satisfy equations (9) since matrices do not commute in general, as we have seen in Chapter V. (The rules for operating with matrices, discussed there, will be used continually in this chapter.) The quantity $\gamma_\mu \partial/\partial x_\mu$ in equation (7) is now both a differential and a matrix operator. It is thus not unreasonable to take ψ to be a column vector, the number of components being equal to the number of rows in γ_μ. γ_μ operating on ψ is then the operation of matrix multiplication. Using Latin letters at the beginning of the alphabet, $a, b, c, \dots$, to label rows and columns of these matrices and vectors (in distinction to $i, j, k, \dots$, used to label spatial components of four-vectors) we write the Dirac equation (7) in the form

$$\sum_b \left(\gamma_\mu^{ab} \frac{\partial}{\partial x_\mu} + M\delta_{ab}\right)\psi^{(b)}(x_\nu) = 0, \tag{7a}$$

or in still more detail as

$$\begin{bmatrix} \gamma_\mu^{11} & \gamma_\mu^{12} \ldots\ldots \\ \gamma_\mu^{21} & \ldots\ldots\ldots\ldots \\ \ldots\ldots & \ldots\ldots\ldots \\ \ldots\ldots & \ldots\ldots\ldots \end{bmatrix} \begin{bmatrix} \dfrac{\partial\psi^{(1)}}{\partial x_\mu} \\ \dfrac{\partial\psi^{(2)}}{\partial x_\mu} \\ \ldots\ldots \\ \ldots\ldots \end{bmatrix} + M \begin{bmatrix} 1 & 0\ldots \\ 0 & 1\ldots \\ \ldots & \ldots\ldots \\ \ldots & \ldots\ldots \end{bmatrix} \begin{bmatrix} \psi^{(1)} \\ \psi^{(2)} \\ \ldots\ldots \\ \ldots\ldots \end{bmatrix} = 0. \qquad (7b)$$

(In both these equations, summation over μ is of course implied.) It is usually most convenient to suppress the indices $a, b, \ldots$, labelling elements of matrices and vectors, as in equations (7) to (9). One must then remember that one is dealing with matrices which are non-commuting. When in doubt and in the course of detailed calculations, it may be advisable to use the more explicit notation employed in equation (7a).

We see that the Dirac equation has one of the properties we desired: the wave function has several components and we shall see that these are connected with spin. For this reason such wave functions are called *spinors**.

We have not so far stated the number of components of these spinor wave functions or the number of rows and columns of the γ-matrices. We shall presently give an explicit representation of the γs, satisfying the anticommutation relations (9), in terms of 4×4 matrices. Furthermore, it can be shown that nothing is gained by considering larger matrices. The corresponding wave functions are four-component spinors.

For many purposes one does not need an explicit matrix representation of the γs but merely the fact that they are quantities satisfying the anticommutation relations (9). One can see that it is possible to get a set of four 4×4 matrices which satisfy the anticommutation relations (9). To show this we note that in virtue of equations (9) any product of an arbitrary number of γs is, apart from sign, equal to one of 16 distinct basic quantities:

$$\left.\begin{array}{lll} \mathbf{I} & (=\text{identity}) & [1] \\ \gamma_\lambda & & [4] \\ \gamma_\lambda\gamma_\mu & (\lambda<\mu) & [6] \\ \gamma_\lambda\gamma_\mu\gamma_\nu & (\lambda<\mu<\nu) & [4] \\ \gamma_1\gamma_2\gamma_3\gamma_4 & & [1] \end{array}\right\} \qquad (10)$$

where $\lambda, \mu, \nu = 1, \ldots 4$ and are subject to the restrictions stated. The numbers in square brackets give the numbers of distinct quantities of each type. Similarly a sum of products of γs can be written as a sum of the above 16 basic quantities with suitable numerical coefficients.

---

* A spinor is rigorously defined by its transformation properties under Lorentz transformations. These properties will be considered in section 49.

(These may of course be complex numbers.) Now a 4 × 4 matrix has 16 elements. It can hence be written as the sum, with suitable coefficients, of 16 linearly independent 4 × 4 matrices. (For these we could, for example, take matrices with zeros everywhere except for one element which is unity, but one can of course equally well take more complicated linearly independent matrices.) It follows that the 16 basic quantities, and in particular the γ_λ themselves, can be represented by 4 × 4 matrices. It is clear from the above that this representation is not unique.

We can write the 16 basic quantities in a different form if we introduce the quantity

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4, \tag{11}$$

which is seen to satisfy the same anticommutation relations as the γ_λ, $(\lambda = 1, \ldots 4)$:

$$\left.\begin{aligned} \gamma_5^2 &= \mathbf{I}, \\ \gamma_\lambda\gamma_5 + \gamma_5\gamma_\lambda &= 0. \end{aligned}\right\} \tag{12}$$

We can then take for our 16 basic quantities:

$$\left.\begin{array}{lll} \mathbf{I} & & [1] \\ \gamma_\lambda & & [4] \\ i\gamma_\lambda\gamma_\mu & (\lambda<\mu) & [6] \\ i\gamma_5\gamma_\lambda & & [4] \\ \gamma_5. & & [1] \end{array}\right\} \tag{13}$$

On account of the anticommutation relations (9) and (12), the quantities (13) agree with those in (10), except for factors $\pm i$. These have been introduced so that the square of each of the basic quantities equals the identity **I**. Furthermore, in the explicit matrix representation which we shall choose, the γ_λ will be Hermitian matrices so that with the above choice of factors all the 16 basic matrices are Hermitian.

The particular 4 × 4 matrix representation which we shall use can be expressed in terms of 2 × 2 submatrices as follows:

$$\gamma_k = \begin{bmatrix} 0 & -i\sigma_k \\ i\sigma_k & 0 \end{bmatrix}, \quad \gamma_4 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{14}$$

and from equation (11) one finds

$$\gamma_5 = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}. \tag{15}$$

Here σ_k are the 2 × 2 Pauli spin matrices [equations (25.31)] and 1 and 0 denote the 2 × 2 unit and null matrices respectively*. These γ-matrices are Hermitian and satisfy the anticommutation relations (9) and (12).

---

* No confusion with the numbers 1 and 0 should arise if one remembers that the γs are 4 × 4 matrices.

In the literature, one often finds, instead of the γ_μ, the related matrices

$$\alpha_k = i\gamma_4\gamma_k, \quad \beta = \gamma_4 \tag{16}$$

whose representation in terms of these of the γs is easily found (Exercise X.2).

We also introduce the generalization of the Pauli spin matrices, which will play the analogous part in the relativistic theory, by

$$\hat{\sigma}_k = \begin{bmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{bmatrix} = i\gamma_4\gamma_5\gamma_k, \quad (k=1, 2, 3) \tag{17}$$

where we have used a 'hat' ($\wedge$) to distinguish the 4×4 matrices. Alternatively we can write these as

$$\hat{\sigma}_k = -i\gamma_l\gamma_m \tag{18}$$

where k, l, m are in cyclic order.

Having treated these matrices in detail, we consider the form the particle density and current take in the relativistic theory. In analogy to the non-relativistic theory (*cf* sections 7 and 26), Dirac defined the particle density at the space–time point x_ν by

$$\left.\begin{aligned} \varrho(x_\nu) &= \psi^*(x_\nu)\psi(x_\nu) \\ &= \sum_{\alpha=1}^{4} |\psi^{(\alpha)}(x_\nu)|^2 \end{aligned}\right\} \tag{19}$$

where ψ^* is the adjoint to ψ, *i.e.* it is the complex conjugate transpose of ψ.

We shall modify our notation somewhat. We shall denote the complex conjugate of A no longer by a bar ($\bar{A}$) but by an asterisk, thus A^*.

We shall also use, instead of the adjoint spinor ψ^*, the related spinor $\bar{\psi}$ defined by

$$\bar{\psi} = \psi^*\gamma_4. \tag{20}$$

The reason for this is that the quantities which have simple transformation properties under Lorentz transformations are naturally written in terms of $\bar{\psi}$ rather than ψ^*. Both ψ^* and $\bar{\psi}$ are of course row vectors and always occur as the left-hand factor in a matrix element, as required by the laws of matrix algebra. One can easily show that if $\psi(x_\nu)$ satisfies the Dirac equation (46.7), then $\bar{\psi}(x_\nu)$ satisfies

$$\frac{\partial\bar{\psi}(x_\nu)}{\partial x_\mu}\gamma_\mu - M\bar{\psi}(x_\nu) = 0 \tag{21}$$

(Exercise X.4).

In terms of $\bar{\psi}$, the density, equation (19), is

$$\varrho(x_\nu) = \bar{\psi}(x_\nu)\gamma_4\psi(x_\nu). \tag{22}$$

Correspondingly, we define a current-density vector $j_k(x_\nu)$ by

$$j_k(x_\nu) = ic\bar{\psi}(x_\nu)\gamma_k\psi(x_\nu). \tag{23}$$

Let us consider the four-vector† $s_\mu(x_\nu)$, defined by

$$s_k = \frac{1}{c}j_k, \quad s_4 = i\varrho \tag{24a}$$

i.e. by

$$s_\mu = i\bar{\psi}\gamma_\mu\psi. \tag{24b}$$

It satisfies the continuity equation

$$\frac{\partial s_\mu}{\partial x_\mu} = 0, \tag{25}$$

(Exercise X.5). This is the justification for interpreting j_k as the current-density. The same problem in the non-relativistic theory was treated in Chapter II (*cf* particularly Exercise II.4).

47. The spin and magnetic moment of the electron

We shall now show that a particle which satisfies the Dirac equation has spin $\frac{1}{2}$. For this purpose we shall show that it is not the orbital angular momentum $\mathbf{l}$ which commutes with the Hamiltonian but the operator*

$$\mathbf{j} = \mathbf{l} + \tfrac{1}{2}\hbar\hat{\boldsymbol{\sigma}} \tag{1}$$

where $\hat{\boldsymbol{\sigma}}$ is the vector with components $\hat{\sigma}_k$.

From equation (46.7) we can at once obtain the Hamiltonian of the Dirac theory. For it can be written

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{2}$$

with

$$H \equiv ic\gamma_4\gamma_k p_k + mc^2\gamma_4 \tag{3}$$

[using equation (46.2) and $M = mc/\hbar$].

Consider the commutator $[l_3, H]$. The second term in equation (3) commutes with l_3. From the first term we obtain, using the results of Exercise IV.13,

$$[l_3, H] = -\hbar c\gamma_4(\gamma_1 p_2 - \gamma_2 p_1), \tag{4}$$

and generally one obtains

$$[\mathbf{l}, H] = -\hbar c\gamma_4(\boldsymbol{\gamma}\wedge\mathbf{p}) \tag{5}$$

where $\boldsymbol{\gamma}$ is the vector γ_k. Thus we see that the orbital angular momentum is not a constant of the motion.

---

† We shall see in section 49 that s_μ is a four-vector.

* Throughout this section $\mathbf{j}$ will denote the operator defined by equation (1). This should not lead to any confusion with the three-current vector which, in the last section, was also denoted by $\mathbf{j}$; equations (46.23) and (46.24).

Consider next $[\hat{\sigma}_3, H]$ which according to equations (46.18) is the same as $[-i\gamma_1\gamma_2, H]$. Since γ_4 commutes with $\gamma_1\gamma_2$, again only the first term in equation (3) need be considered. Now from the anti-commutation relations (46.9) one easily proves that

$$\left.\begin{aligned}[\gamma_1\gamma_2, \gamma_4\gamma_1] &= -2\gamma_4\gamma_2\\ [\gamma_1\gamma_2, \gamma_4\gamma_2] &= +2\gamma_4\gamma_1\\ [\gamma_1\gamma_2, \gamma_4\gamma_3] &= 0\end{aligned}\right\} \tag{6}$$

so that

$$[-i\gamma_1\gamma_2, H] = 2c\gamma_4(\gamma_1 p_2 - \gamma_2 p_1) \tag{7}$$

and generally

$$[\hat{\boldsymbol{\sigma}}, H] = 2c\gamma_4(\boldsymbol{\gamma}\wedge\mathbf{p}). \tag{8}$$

From equations (1), (5) and (8) it follows that

$$[\mathbf{j}, H] = [\mathbf{l} + \tfrac{1}{2}\hbar\hat{\boldsymbol{\sigma}}, H] = 0. \tag{9}$$

This justifies our interpretation of **j** as the total angular momentum of the particle with the resulting conclusions about its spin. In section 48, dealing with solutions of the Dirac equation of a free particle we shall come back to the question of spin. The proof which we have here given also applies in the more general case of a Dirac particle in a spherically symmetric external field (for example in the hydrogen atom). Such a potential leads to an additional term in the Hamiltonian which commutes with all rotations and hence with the l_s (*cf* Sections 41 and 42). Since it depends only on the spatial variables and not on the spin, it also commutes with $\hat{\boldsymbol{\sigma}}$, and hence with the total angular momentum **j**, as required from general considerations. In section 49 it will be shown that the operator $\hat{\boldsymbol{\sigma}}$ is a pseudo-vector as required for an angular momentum [*cf* equations (41.8) and (41.9)].

We next investigate the magnetic moment, associated with this spin of the electron. For that purpose we consider the effect of a magnetic field **H**, constant in time. Such a field is described by a time-independent vector potential **A** (*i.e.* by a four-vector A_μ with $A_4 = 0$), and is given by (*cf* Abraham–Becker[6])

$$\mathbf{H} = \operatorname{curl} \mathbf{A}. \tag{10}$$

It can be shown quite generally that the effect of an electromagnetic field specified by a four-vector $A_\mu(x_\nu)$ on a particle of charge $(-e)$ is correctly described by replacing the four-momentum p_μ in the equations of motion by the four-vector*

$$\left(p_\mu + \frac{e}{c}A_\mu\right). \tag{11}$$

---

* This result is obtained from the relativistic equation of motion of a charged particle in an electromagnetic field, as given by Lorentz. It is shown that this equation can be derived from a Lagrangian by means of the usual Lagrangian or Hamiltonian formalisms provided that the momentum p_μ is replaced by the four-vector $(p_\mu + eA_\mu/c)$ throughout. For details, *cf* Mott and Sneddon[10], §52.

Thus, the Dirac equation of an electron in the field A_μ becomes, on account of (46.2),

$$\gamma_\mu\left(\frac{\partial}{\partial x_\mu}+\frac{ie}{\hbar c}A_\mu\right)\psi+M\psi=0. \tag{12}$$

To interpret this, we again convert this to a second-order equation by operating on equation (12) with

$$\left[\gamma_\lambda\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)-M\right],$$

giving

$$\left[\gamma_\lambda\gamma_\mu\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)\left(\frac{\partial}{\partial x_\mu}+\frac{ie}{\hbar c}A_\mu\right)-M^2\right]\ \psi=0. \tag{13}$$

The terms with $\lambda=\mu$ give (since $\gamma_\lambda{}^2=1$)

$$\sum_\lambda\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)^2\psi. \tag{14a}$$

The terms with $\lambda\neq\mu$ give [on account of equations (46.9)]

$$\tfrac{1}{2}\gamma_\lambda\gamma_\mu\left\{\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)\left(\frac{\partial}{\partial x_\mu}+\frac{ie}{\hbar c}A_\mu\right)-\left(\frac{\partial}{\partial x_\mu}+\frac{ie}{\hbar c}A_\mu\right)\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)\right\}\psi$$

$$=\tfrac{1}{2}\gamma_\lambda\gamma_\mu\left(\frac{ie}{\hbar c}\right)\left(\frac{\partial A_\mu}{\partial x_\lambda}-\frac{\partial A_\lambda}{\partial x_\mu}\right)\psi. \tag{14b}$$

Combining these results, equation (13) becomes

$$\left[\sum_\lambda\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)^2+\frac{ie}{2\hbar c}\gamma_\lambda\gamma_\mu\left(\frac{\partial A_\mu}{\partial x_\lambda}-\frac{\partial A_\lambda}{\partial x_\mu}\right)-M^2\right]\psi=0. \tag{15}$$

From the definition of curl (Appendix A), equations (10) and (46.18), one obtains

$$\left[\sum_\lambda\left(\frac{\partial}{\partial x_\lambda}+\frac{ie}{\hbar c}A_\lambda\right)^2-\frac{e}{\hbar c}\hat{\boldsymbol{\sigma}}\mathbf{H}-M^2\right]\psi=0, \tag{16}$$

remembering that we are considering a static magnetic field. We write this equation as

$$\left(\frac{\partial}{\partial x_4}-M\right)\left(\frac{\partial}{\partial x_4}+M\right)\psi=\left[-\sum_k\left(\frac{\partial}{\partial x_k}+\frac{ie}{\hbar c}A_k\right)^2+\frac{e}{\hbar c}\hat{\boldsymbol{\sigma}}\mathbf{H}\right]\psi \tag{17}$$

and take the non-relativistic limit, by considering a particle whose velocity is small compared to the velocity of light. Its energy will then be very nearly equal to its rest energy mc^2. Since for an eigenstate of energy E,

$$i\hbar\frac{\partial\psi}{\partial t}=E\psi,$$

it follows that in the *first* factor on the left-hand side of equation (17) we can put $i\hbar\frac{\partial}{\partial t}=mc^2$ *i.e.*

$$\frac{\partial}{\partial x_4}=-M. \tag{18}$$

Equation (17) then gives the 'Schrödinger equation'

$$i\hbar\frac{\partial\psi}{\partial t}=\left[mc^2-\frac{\hbar^2}{2m}\sum_{k=1}^{3}\left(\frac{\partial}{\partial x_k}+\frac{ie}{\hbar c}A_k\right)^2+\frac{e\hbar}{2mc}\hat{\boldsymbol{\sigma}}\mathbf{H}\right]\psi. \tag{19}$$

Since the potential energy of a magnet of magnetic moment **M** in a magnetic field **H** is (−**MH**), the last term in equation (19) shows that associated with the electron, as described by the Dirac equation, is a magnetic moment

$$\frac{-e}{mc}(\tfrac{1}{2}\hbar\hat{\boldsymbol{\sigma}}), \tag{20}$$

as required by experiment. The correct prediction of this magnetic moment was one of the first and most remarkable successes of the Dirac theory of the electron. Since then it has had overwhelming detailed confirmation in many branches of physics; for example in quantum electrodynamics where it leads to an excellent description of the interaction of electrons with the electromagnetic field.

48. Plane wave solutions of the Dirac equation

Let us obtain solutions of the Dirac equation of a free particle, equation (46.7). We expect such solutions to exist, representing plane waves corresponding to states of definite momentum **p**. From the relativistic Hamiltonian (46.1) it follows that for a particle of momentum **p** the energy E can assume two values

$$E=\pm E_p \tag{1}$$

where

$$E_p=+(m^2c^4+c^2p^2)^{\frac{1}{2}}. \tag{2}$$

The existence of these negative energy states leads to difficulties in a naïve interpretation of the theory. We shall return to this point later and, to begin with, find the explicit solutions.

For a given value of **p** and E_p, the Dirac equation admits four linearly independent solutions, which we can take to be

$$\psi_r=\left\{\begin{array}{ll} u_r(p_\nu)e^{ip_\nu x_\nu/\hbar} & (r=1,2) \\ u_r(p_\nu)e^{-ip_\nu x_\nu/\hbar} & (r=3,4) \end{array}\right\} \tag{3}$$

where $p_\nu=(\mathbf{p}, iE_p/c)$. The first two of these solutions $(r=1, 2)$ correspond to electron states with momentum **p** and energy E_p, the last two

($r=3, 4$) to momentum $-\mathbf{p}$ and energy $-E_p$. That there exist exactly four solutions follows from the fact that the Dirac equation really represents a system of four simultaneous linear equations for the four components of the wave function. In equation (3), the u_r are constant four-component spinor wave-functions (*i.e.* for each fixed r, the u_r has four components $u_r^{(a)}$, $a=1, \ldots 4$). To find these constant spinors, we substitute equations (3) into the Dirac equation and solve the resulting set of linear equations for the u_r. For this purpose we must of course use a particular matrix representation for the γ-matrices, for example that defined in section 46 which we shall use throughout. One then finds:

$$u_1(p_\nu)=\left(\frac{E_p+mc^2}{2mc^2}\right)^{\frac{1}{2}}\begin{bmatrix}1\\0\\cp_3/(E_p+mc^2)\\c(p_1+ip_2)/(E_p+mc^2)\end{bmatrix}, \tag{4a}$$

$$u_2(p_\nu)=\left(\frac{E_p+mc^2}{2mc^2}\right)^{\frac{1}{2}}\begin{bmatrix}0\\1\\c(p_1-ip_2)/(E_p+mc^2)\\-cp_3/(E_p+mc^2)\end{bmatrix}, \tag{4b}$$

$$u_3(p_\nu)=\left(\frac{E_p+mc^2}{2mc^2}\right)^{\frac{1}{2}}\begin{bmatrix}cp_3/(E_p+mc^2)\\c(p_1+ip_2)/(E_p+mc^2)\\1\\0\end{bmatrix}, \tag{4c}$$

$$u_4(p_\nu)=\left(\frac{E_p+mc^2}{2mc^2}\right)^{\frac{1}{2}}\begin{bmatrix}c(p_1-ip_2)/(E_p+mc^2)\\-cp_3/(E_p+mc^2)\\0\\1\end{bmatrix}. \tag{4d}$$

The multiplicative factor $[(E_p+mc^2)/2mc^2]^{\frac{1}{2}}$ is of course not determined by the Dirac equation which is homogeneous in the unknowns. We have chosen it to give a convenient normalization. We shall not here deal with the orthogonality and normalization properties of the plane wave solutions (3) and (4). The most important properties, demonstrating the expediency of our choice, are left as exercises to the reader (Exercises X.6–7).

The Dirac equation and its solutions become particularly simple in the rest frame of the electron: $\mathbf{p}=0$. Substituting equations (3) into the Dirac equation for this case, one obtains for the positive energy solutions

$$(\gamma_4-1)u_r=0, \quad (r=1, 2),$$

and for the negative energy solutions

$$(\gamma_4+1)u_r=0, \quad (r=3, 4).$$

In our particular matrix representation, γ_4 is given by equation (46.14). It follows at once that the eigenvectors u_r are given by

$$u_1=\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\quad u_2=\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\quad u_3=\begin{bmatrix}0\\0\\1\\0\end{bmatrix},\quad u_4=\begin{bmatrix}0\\0\\0\\1\end{bmatrix}. \tag{5}$$

This agrees with equations (4a–d).

We next consider the spin of the electron in its rest system. We see from the definition of the 4×4 spin matrix $\hat{\boldsymbol{\sigma}}$, equation (46.17), that

$$\left.\begin{array}{ll}\hat{\sigma}_3u_1=u_1 & \hat{\sigma}_3u_3=u_3\\ \hat{\sigma}_3u_2=-u_2 & \hat{\sigma}_3u_4=-u_4\end{array}\right\} \tag{6}$$

i.e. in the rest system we can interpret u_1 and u_3 as states of the electron with spin pointing in the positive z-direction (one of positive, the other of negative energy); similarly u_2 and u_4 correspond to states with the spin pointing in the negative z-direction. This interpretation can still be approximately maintained in the non-relativistic case of small velocities. In general it breaks down: the u_r are not eigenfunctions of $\hat{\sigma}_z$. This is not surprising, since we have seen (section 47) that the spin is not a constant of the motion of the Dirac Hamiltonian. However, the component of the spin vector $\hat{\boldsymbol{\sigma}}$ in the direction of the momentum $\mathbf{p}$ is a constant of the motion (Exercise X.8).

We consider the non-relativistic case of particles whose velocity v is small compared to the velocity of light: $v\ll c$. Then $E_p\simeq mc^2$ and $p_k\simeq mv_k$, so that terms of the form $cp_k/(E_p+mc^2)$ are of order v/c and hence small in the non-relativistic limit. Hence in the positive energy solutions, u_1 and u_2, the third and fourth components are small, and in the negative energy solutions, u_3 and u_4, it is the first and second components that are small in the non-relativistic limit. These conclusions are of course consistent with the interpretation of equations (5) as representing the eigenvectors in the rest system of the electron.

In practice one often requires the non-relativistic limit of such matrix elements as $(u_r, \gamma_\mu u_s)$. To deal with this question, we remember that the positive energy solutions u_1 and u_2 (we shall briefly write either of them u_+) decompose into two large components (1 and 2) and two small components (3 and 4), so that we can write

$$u_+=\begin{bmatrix}u_+^L\\u_+^S\end{bmatrix}. \tag{7a}$$

Here u_+^L itself stands for the two-component vector of the large components, *etc.* Similarly we write for either of the negative energy solutions u_3 or u_4,

$$u_-=\begin{bmatrix}u_-^S\\u_-^L\end{bmatrix}. \tag{7b}$$

Remembering that the u^S are of order $(v/c)u^L$, and using the explicit matrix representations of the γ-matrices (whose elements are 0, ± 1, $\pm i$) one can easily give the order of magnitude of matrix elements. As an example, we calculate

$$\begin{aligned}(u_+, \gamma_4 u_+) &= [u_+^{L\star}, u_+^{S\star}] \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} u_+^L \\ u_+^S \end{bmatrix} \\ &= u_+^{L\star} u_+^L - u_+^{S\star} u_+^S \\ &\simeq u_+^{L\star} u_+^L \\ &= (u_+^L, u_+^L), \end{aligned} \tag{8}$$

neglecting a term of order $(v/c)^2$. We have thus reduced this matrix element to its non-relativistic form in terms of two-component wavefunctions.

An essential difference in such matrix elements arises between diagonal and off-diagonal operators, *i.e.* matrices of the types

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & A \\ B & 0 \end{bmatrix}$$

respectively. (Examples of diagonal matrices are γ_4 and $\hat{\sigma}_k$; of off-diagonal matrices: γ_k and γ_5.) For they mix up large and small components differently. As an example of an off-diagonal matrix, consider $(u_+, \gamma_5 u_+)$:

$$\begin{aligned}(u_+, \gamma_5 u_+) &= [u_+^{L\star}, u_+^{S\star}] \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} u_+^L \\ u_+^S \end{bmatrix} \\ &= -u_+^{L\star} u_+^S - u_+^{S\star} u_+^L. \end{aligned} \tag{9}$$

Here large and small components are mixed up and so this matriz element is of order v/c compared to (u_+, u_+). But of course γ_5 gives large matrix elements between positive and negative energy states; *e.g.*

$$\begin{aligned}(u_+, \gamma_5 u_-) &= [u_+^{L\star}, u_+^{S\star}] \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} u_-^S \\ u_-^L \end{bmatrix} \\ &= -u_+^{L\star} u_-^L - u_+^{S\star} u_-^S \\ &\simeq -u_+^{L\star} u_-^L, \end{aligned} \tag{10}$$

neglecting terms of order $(v/c)^2$. These examples should suffice to enable the reader to deal with any case. It should again be stressed that the explicit solutions and detailed deduction from them, made in this section, depend on the particular matrix representation which we have chosen for the γ_μ.

We come back to the physical interpretation of the negative energy states. One is here faced with the difficulty that quantum jumps from the positive to the negative energy states are possible; so we cannot ignore the negative energy solutions, as is done in classical theory.

To begin with, we must admit that these difficulties cannot be properly resolved within the framework of a one-particle theory such as developed here. One requires a many-particle theory in which electrons can be created and absorbed, *i.e.* a field theory, analogous to Maxwell's theory of the electromagnetic field which is a field theory for photons. As we shall not treat field theory in this book we can only give a very qualitative discussion of how the difficulties associated with the negative energy states are resolved.

For this purpose, Dirac postulated that the negative energy states are ordinarily all occupied. (Remember, we are dealing with electrons so that, according to the Pauli principle, there can at most be one particle in each state.) This state with all negative energy levels filled is then to be thought of as the vacuum and the negative energy sea of electrons are not to produce any physical effects.

An unoccupied negative energy state, *i.e.* a *hole* in the sea of negative energy electrons, with energy $-E_p$ and momentum $-\mathbf{p}$ can then be interpreted as a *positron* of energy E_p and momentum $\mathbf{p}$. The description of electrons and positrons which we have here given appears very unsymmetrical in positively and negatively charged particles. In fact one can show that the field theory is completely symmetrical in electrons and positrons, the apparent asymmetry being due to the particular way we introduced positrons starting from the single particle Dirac equation for electrons. The theory developed along these lines is remarkably successful. The existence of the positron, predicted by Dirac from the above theory, has been confirmed and the theory affords a quantitatively correct description of such phenomena as Bremsstrahlung and Compton scattering, as well as pair creation and pair annihilation. (Pair creation is the process when an electron from the negative energy sea is excited into a positive energy state leaving an empty hole, a positron, the excitation energy being provided by a high-energy photon, *i.e.* by a gamma-ray. Pair annihilation is the inverse process.)

49. Relativistic invariance

We now wish to consider certain aspects of the relativistic invariance of the Dirac theory which are especially important in practice.

We begin by stating briefly the basic properties of Lorentz transformations. Consider a linear homogeneous transformation from a coordinate system x_ν to a coordinate system x_ν', defined by

$$x_\mu' = a_{\mu\nu} x_\nu. \tag{1}$$

This transformation is a homogeneous Lorentz transformation if it leaves the lengths of vectors invariant, *i.e.*

$$\sum_{\nu=1}^{4} x_\nu'^2 = \sum_{\nu=1}^{4} x_\nu^2. \tag{2}$$

This is the condition that the velocity of light should have the same value c in all Lorentz frames. Substitution of equation (1) into (2) gives a condition on the transformation matrix $a_{\mu\nu}$:

$$a_{\mu\nu} a_{\mu\sigma} = \delta_{\nu\sigma}. \tag{3}$$

Using this equation one finds at once the inverse of equation (1):

$$x_\nu = a_{\mu\nu} x_\mu', \tag{1a}$$

and substituting equation (1a) into (2), one finds, analogously to (3):

$$a_{\lambda\nu} a_{\mu\nu} = \delta_{\lambda\mu}. \tag{3a}$$

Equations (3a) are however *not* additional new conditions on the matrix $a_{\mu\nu}$ but follow from equations (3). The determinant of a Lorentz transformation has the value ±1. For, from equation (3),

$$\begin{aligned}(|a_{\mu\alpha}|)^2 = |a_{\mu\alpha}||a_{\nu\beta}| &= |a^T_{\alpha\mu}||a_{\nu\beta}| \\ &= |a^T_{\alpha\nu} a_{\nu\beta}| \\ &= |\delta_{\alpha\beta}| = 1,\end{aligned}$$

where the index T denotes the transpose. Hence

$$|a_{\nu\alpha}| = \pm 1. \tag{4}$$

We now restrict the transformations to the *proper* Lorentz transformations by confining ourselves to transformations with determinant +1:

$$|a_{\lambda\mu}| = 1, \tag{5}$$

and in addition we require that

$$a_{44} > 0. \tag{6}$$

This means that we do not consider transformations which reverse the sense of time, interchanging future and past light cones.

Of the improper Lorentz transformations (which are defined by having a determinant equal to −1) we shall consider spatial inversion, defined by the diagonal matrix

$$[a_{\mu\nu}] = \begin{bmatrix} -1 & & & \\ & -1 & & \\ & & -1 & \\ & & & +1 \end{bmatrix}, \tag{7}$$

agreeing with our earlier definition (p. 39 and sections 39 and 41):

$$x_k \rightarrow x_k' = -x_k, \quad x_4' = x_4. \tag{8}$$

Thus inversion is the same as going from a right-handed to a left-handed spatial coordinate system.

The proper Lorentz transformations include as a particular case the identity

$$a_{\mu\nu} = \delta_{\mu\nu} \tag{9}$$

which simply leaves the coordinate system unchanged. One easily shows that the proper Lorentz transformations form a group, in the sense in which we used this term in Chapter IX (Exercise X.9). Hence one need for many purposes only consider the infinitesimal Lorentz transformations.

Lastly, we must mention the reality character of the matrix elements $a_{\mu\nu}$. From our original definition (Section 46) the x_k are real and x_4 is pure imaginary; we require this of course in all coordinate systems. One easily sees that this is ensured if we have

$$\left.\begin{array}{ll} a_{kl},\, a_{44} & \text{real} \\ a_{k4},\, a_{4k} & \text{pure imaginary} \end{array}\right\}. \tag{10}$$

This property is also preserved by two successive Lorentz transformations.

We now wish to study the properties of the Dirac equation which follow from the fact that we are trying to formulate a relativistically invariant theory. We would like to remind the reader of what we mean by relativistic invariance, using Maxwell's wave equations for the four-potential $A_\mu(x_\nu)$ as an example. An observer O, using a coordinate system (x_ν), will write these

$$\frac{\partial^2}{\partial x_\lambda \partial x_\lambda} A_\mu(x_\nu) = e s_\mu(x_\nu),$$

where s_μ is the classical charge-current density. Consider a second observer O' using a coordinate system (x_ν') related to the system (x_ν) by a Lorentz transformation. The principle of special relativity then requires that there should be nothing to choose between the two coordinate systems: the observer O should in no way have a preferred position with respect to O' or *vice versa*. This means that O' must write Maxwell's equations as

$$\frac{\partial^2}{\partial x_\lambda' \partial x_\lambda'} A_\mu'(x_\nu') = e s_\mu'(x_\nu'),$$

i.e. the equations must be *invariant in form* under a Lorentz transformation, just as in classical mechanics Newton's equations of motion are invariant in form under a Galilean transformation. This invariance in form follows at once. From equations (1) and (3) $\partial^2/\partial x_\lambda \partial x_\lambda$ is invariant under Lorentz transformations and A_μ and s_μ are four-vectors and thus transform in the same way under Lorentz transformations.

After these preliminaries we discuss the relativistic invariance of the Dirac equation

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu} + M\right)\psi(x_\nu) = 0. \tag{11}$$

We shall take the γ-matrices as constant, *i.e* the same in any Lorentz frame. Consider a particular Lorentz transformation $a_{\mu\nu}$, equation (1). This will induce a linear transformation in the four-dimensional spinor space (just as in section 42 we could think of three-dimensional rotations inducing a transformation in spin-space) so that the same state of the system will be represented by a new spinor $\psi'(x_\nu')$. The Dirac equation is invariant under Lorentz transformations if equation (11) transforms into

$$\left(\gamma_\mu \frac{\partial}{\partial x_\mu'} + M\right)\psi'(x_\nu') = 0. \tag{12}$$

This is the case provided that, for a given Lorentz transformation, $\psi'(x_\nu')$ can be written in the form

$$\psi'(x_\nu') = S\psi(x_\nu) \tag{13}$$

where S is a constant 4×4 matrix, determined by the Lorentz transformation $a_{\mu\nu}$, which possesses an inverse S^{-1}. Using equations (1) and (13), the Dirac equation (11) becomes

$$\gamma_\mu \frac{\partial}{\partial x_\lambda'} a_{\lambda\mu} S^{-1}\psi'(x_\nu') + MS^{-1}\psi'(x_\nu') = 0,$$

i.e. premultiplying by S:

$$a_{\lambda\mu} S\gamma_\mu S^{-1} \frac{\partial}{\partial x_\lambda'}\psi'(x_\nu') + M\psi'(x_\nu') = 0. \tag{14}$$

This equation is identical with equation (12) if

$$a_{\lambda\mu} S\gamma_\mu S^{-1} = \gamma_\lambda \tag{15a}$$

or, rewriting this, if

$$S^{-1}\gamma_\lambda S = a_{\lambda\mu}\gamma_\mu. \tag{15b}$$

In section 46 above we defined the charge-current density s_μ. We require s_μ to be a four-vector, so that the continuity equation (46.25) permits the usual invariant interpretation. This means that it must transform according to

$$s_\mu(x_\lambda) \rightarrow s_\nu'(x_\lambda') = a_{\nu\mu} s_\mu(x_\lambda),$$

i.e.

$$\bar{\psi}'\gamma_\nu\psi' = a_{\nu\mu}\bar{\psi}\gamma_\mu\psi.$$

On account of equations (13) and 15b) we can write this equation as

$$a_{\nu\mu}\bar{\psi}' S\gamma_\mu\psi = a_{\nu\mu}\bar{\psi}\gamma_\mu\psi.$$

Hence $\bar{\psi}(x_\lambda)$ must transform according to

$$\bar{\psi}'(x_\lambda') = \bar{\psi}(x_\lambda)S^{-1}. \tag{16}$$

But from equation (13) itself one finds directly that

$$\begin{aligned}\bar{\psi}'(x_\lambda') = \psi'^{\star}(x_\lambda')\gamma_4 &= \psi^{\star}(x_\lambda)S^{\star}\gamma_4 \\ &= \bar{\psi}(x_\lambda)\gamma_4 S^{\star}\gamma_4.\end{aligned}$$

Comparison with equation (16) shows that, for consistency, we must require S to satisfy the condition

$$\gamma_4 S^{\star}\gamma_4 = S^{-1}. \tag{17}$$

To complete the proof of the Lorentz invariance of the Dirac equation, it must be shown that for every Lorentz transformation $a_{\mu\nu}$ a matrix S with the required properties (15) and (17) exists. Since the Lorentz transformations form a group, it suffices to construct this matrix for the infinitesimal Lorentz transformations from which any finite Lorentz transformation can then be generated. We shall omit the proof of this existence theorem (*cf*, for example, reference 17; or in English: S. S. Schweber, H. A. Bethe and F. de Hoffmann, *Mesons and Fields*, volume 1). For our purposes it suffices to know that a matrix S with these properties, equations (15) and (17), exists for every Lorentz transformation.

From the ψ and $\bar{\psi}$ and the γ-matrices one can construct five quantities which have definite transformation properties under Lorentz transformations; they are scalars, vectors, *etc.* These five *covariants*—as quantities with definite transformation properties are called—play a very important part in relativistic quantum mechanics, for example in the theory of β-decay (Bethe and Morrison[19], Blatt and Weisskopf[20]). These five covariants are closely related to the five combinations of γ-matrices given in equations (46.10) or (46.13). We define them as:

(*i*) scalar or invariant (1 component):

$$I = \bar{\psi}\psi. \tag{18a}$$

(*ii*) four-vector (4 components):

$$s_\mu = i\bar{\psi}\gamma_\mu\psi. \tag{18b}$$

We have denoted this by s_μ as it is identical with the charge-current density vector considered in section 46, equations (46.22–25).

(*iii*) antisymmetric tensor of the second rank or six-vector (6 components):

$$\left.\begin{aligned} M_{\mu\nu} &= i\bar{\psi}\gamma_\mu\gamma_\nu\psi, \quad \mu\neq\nu, \\ &= \quad 0 \quad , \quad \mu=\nu, \end{aligned}\right\} \tag{18c}$$

and one sees at once that $M_{\mu\nu} = -M_{\nu\mu}$.

(*iv*) pseudo-vector or axial vector (4 components):

$$K_\mu = i\bar{\psi}\gamma_5\gamma_\mu\psi. \tag{18d}$$

(*v*) pseudo-scalar (1 component):

$$\mathcal{N} = i\bar{\psi}\gamma_5\psi. \tag{18e}$$

The factors i have been included in these definitions so that the different components should have the reality character required of them. We see, for example, from equation (18b) that s_μ has real spatial components, whereas the time-like component s_4 is purely imaginary [compare equation (46.24a)]. Similarly in equation (18c) the M_{kl} are real, the $M_{k4} = -M_{4k}$ purely imaginary, as required for a six-vector.

It is now easy to demonstrate the transformation properties of the above quantities, using equations (13), (16) and (15). We have already seen that these equations imply that s_μ, equation (18b), transforms like a four-vector:

$$s_\mu' = a_{\mu\nu}s_\nu. \tag{19b}$$

We consider two of the other cases as typical examples:

(*i*) scalar: from equation (18a)

$$I' = \bar{\psi}'\psi' = \bar{\psi}S^{-1}S\psi = \bar{\psi}\psi = I. \tag{19a}$$

(*ii*) pseudo-scalar: from equation (18e),

$$\begin{aligned}\mathcal{N}' = i\bar{\psi}'\gamma_5\psi' &= i\bar{\psi}S^{-1}\gamma_5S\psi \\ &= i\bar{\psi}S^{-1}\gamma_1\gamma_2\gamma_3\gamma_4S\psi \\ &= i\bar{\psi}S^{-1}\gamma_1SS^{-1}\gamma_2\ldots S\psi \\ &= a_{1\lambda}a_{2\mu}a_{3\nu}a_{4\sigma}i\bar{\psi}\gamma_\lambda\gamma_\mu\gamma_\nu\gamma_\sigma\psi.\end{aligned} \tag{20}$$

In this four-fold sum over $\lambda, \mu, \nu, \sigma$ only terms with all four indices different from each other contribute. [This follows, since for any fixed ν and σ say, the sum over the four terms $\lambda = \mu$ vanishes on account of the anticommutation relations (46.9) and the orthogonality relation (3a).] We define $\varepsilon_{\lambda\mu\nu\sigma}$ to be zero if the indices $\lambda, \mu, \nu, \sigma$ are not all different, to be equal to $+1$ if they are in the order 1234 or an even permutation of this order and to be equal to -1 if they are an odd permutation of this order. Equation (20) can then be written in the form

$$\mathcal{N}' = a_{1\lambda}a_{2\mu}a_{3\nu}a_{4\sigma}\varepsilon_{\lambda\mu\nu\sigma}i\bar{\psi}\gamma_1\gamma_2\gamma_3\gamma_4\psi. \tag{21}$$

The ε-symbol in this equation originates from the rearrangement of $\gamma_\lambda\gamma_\mu\gamma_\nu\gamma_\sigma$ in equation (20) in the order $\gamma_1\gamma_2\gamma_3\gamma_4$ using the anticommutation relations (46.9). Now one can easily see that the determinant $|a_{\mu\nu}|$ can be expanded in the form

$$|a_{\mu\nu}| = \varepsilon_{\lambda\mu\nu\sigma}a_{1\lambda}a_{2\mu}a_{3\nu}a_{4\sigma}, \tag{22}$$

so that equation (21) finally gives

$$\mathcal{N}' = |a_{\mu\nu}|\mathcal{N}. \tag{19e}$$

Hence it follows that under proper Lorentz transformations ($|a_{\mu\nu}| = +1$) $\mathcal{N}$ is invariant, whereas under improper Lorentz transformations ($|a_{\mu\nu}| = -1$), in particular under inversion, it changes sign. These are exactly the transformation properties which define a pseudo-scalar.

The same method is easily applied to the remaining two covariants (18c) and (18d). One finds

$$M_{\mu\nu}' = a_{\mu\sigma}a_{\nu\tau}M_{\sigma\tau} \tag{19c}$$

and

$$K_{\mu}' = |a_{\sigma\tau}|a_{\mu\nu}K_{\nu}. \tag{19d}$$

which again are the characteristic transformation properties of tensors of the second rank and of pseudo-vectors respectively. In particular, under inversion, equation (7), $K_i \to K_i' = K_i$, whereas a vector changes sign, *e.g.* $\mathbf{r} \to \mathbf{r}' = -\mathbf{r}$. We note in particular, as stated earlier (section 47), that the spin $\hat{\sigma}_k$ transforms like a pseudo-vector. On account of equation (46.17) we can write $\psi^*\hat{\sigma}_k\psi$ as

$$\psi^*\hat{\sigma}_k\psi = i\bar{\psi}\gamma_5\gamma_k\psi, \tag{23}$$

which is the spatial part of the pseudo-vector K_μ, equation (18d).

The importance of these covariants is due to the fact that from them one builds up relativistically invariant expressions (analogous to the scalar product $x_\mu y_\mu$) which represent the interactions between particles or fields. The requirement of invariance places a very severe restriction on the possible forms of interaction. The interested reader is referred to the theory of β-decay which affords a particularly good example of these matters.

EXERCISES X

X.1. Show that the γ-matrices defined in equations (46.14–15) satisfy the commutation relations (46.9) and (46.12).

X.2. Derive the matrix representations of α_k, equation (46.16), from (46.14).

X.3. Derive equations (46.18) from (46.17).

X.4. Prove equation (46.21).

X.5. Prove the continuity equation (46.25).

X.6. For the spinors u_r, defined in equations (48.4), prove that:

$$(i)\quad (u_1, u_2) = (u_3, u_4) = 0$$

but all other inner products $(u_r, u_s) \neq 0$.

$$(ii)\quad \bar{u}_r u_s = 0, \text{ if } r \neq s.$$

X.7. For the complete plane wave solutions ψ_r, equations (48.3), show that

(*i*) $(\psi_r, \psi_s) = 0$ if $r \neq s$;

(*ii*) the ψ_r are normalized to E_p/mc^2 particles per unit volume;

(*iii*) this normalization is invariant under Lorentz transformations.

X.8. Prove that the component of the spin vector $\hat{\boldsymbol{\sigma}}$ in the direction of the momentum $\mathbf{p}$, *i.e.* $(\hat{\boldsymbol{\sigma}}\mathbf{p})$, is a constant of the motion for the Dirac equation.

X.9. Prove the group property of the proper Lorentz transformations.

HINTS FOR SOLVING THE EXERCISES

CHAPTER I

I.1. THE fundamental domain being $-L<x<L$, the inner product of two functions $f(x)$ and $g(x)$ is defined by

$$(g(x), f(x)) = \int_{-L}^{L} \overline{g(x)} f(x) dx.$$

To prove that the set is normed and orthogonal one need only show that

$$(\varphi_m, \varphi_n) = \delta_{mn}, \quad (\psi_m, \psi_n) = \delta_{mn}, \quad (\psi_m, \varphi_n) = 0 \qquad \text{(H 1)}$$

for all m and n.

I.2. To obtain the a_n and b_n, we successively form the inner product of equation (I.2) with ψ_m and φ_m using equations (H 1). For example

$$(\varphi_m(x), f(x)) = b_m$$

i.e

$$b_m = \frac{1}{\sqrt{L}} \int_{-L}^{L} f(x) \sin \frac{m\pi}{L} x dx \quad (m=1, 2, \ldots),$$

and similarly

$$a_0 = \frac{1}{\sqrt{2L}} \int_{-L}^{L} f(x) dx$$

$$a_m = \frac{1}{\sqrt{L}} \int_{-L}^{L} f(x) \cos \frac{m\pi}{L} x dx \quad (m=1, 2, \ldots).$$

To show that the φ_n, ψ_n $(n=0, 1, 2, \ldots)$ and the χ_n $(n=0, \pm 1, \pm 2, \ldots)$ span the same space, if $L=\pi$, we have that

(*i*) $\psi_0 = \chi_0$.

(*ii*) For any positive integer n, the manifold spanned by φ_n, ψ_n is the same as that spanned by χ_n, χ_{-n}, in fact

$$\cos nx = \tfrac{1}{2}(e^{inx} + e^{-inx})$$

$$\sin nx = \tfrac{1}{2i}(e^{inx} - e^{-inx})$$

i.e.

$$\psi_n = \frac{1}{\sqrt{2}}(\chi_n + \chi_{-n})$$

$$\varphi_n = \frac{1}{i\sqrt{2}}(\chi_n - \chi_{-n}).$$

Hence any vector of the form $c_n\varphi_n + d_n\psi_n$ can be written in the form $C_n\chi_n + D_n\chi_{-n}$.

I.3. If $f(x)$ is odd, *i.e.* $f(-x)=-f(x)$, then since $\sin(-\lambda x) = -\sin(\lambda x)$ and $\cos(-\lambda x)=\cos(\lambda x)$,

$$f(-x)\cos\left[-\frac{m\pi}{L}x\right] = -f(x)\cos\left[\frac{m\pi}{L}x\right]$$

$$f(-x)\sin\left[-\frac{m\pi}{L}x\right] = f(x)\sin\left[\frac{m\pi}{L}x\right]$$

and therefore

$$\left.\begin{aligned} &a_0=a_m=0,\\ &b_m=\frac{2}{\sqrt{L}}\int_0^L f(x)\sin\frac{m\pi}{L}x\,dx, \end{aligned}\right\} \quad (m=1, 2, 3, \dots),$$

in agreement with the analysis in section 2. If $f(x)$ is even, *i.e.* $f(-x)=f(x)$, the general Fourier series similarly reduces to a cosine series. We see that we can think of the normed orthogonal sets $[\psi_0, \psi_1, \psi_2, \dots]$ and $[\varphi_1, \varphi_2, \dots]$ as defining mutually orthogonal subspaces of the total space spanned by the ψ_n and the φ_n together.

I.4. From Exercise I.2 we have

$$a_0=\frac{1}{\sqrt{2\pi}}\int_0^\pi x\,dx=\frac{1}{\sqrt{2\pi}}\frac{\pi^2}{2},$$

$$a_n=\frac{1}{\sqrt{\pi}}\int_0^\pi x\cos nx\,dx=\frac{1}{\sqrt{\pi}n^2}(\cos n\pi-1)=\frac{1}{\sqrt{\pi}n^2}[(-1)^n-1],$$

$$b_n=\frac{1}{\sqrt{\pi}}\int_0^\pi x\sin nx\,dx=\frac{-\sqrt{\pi}}{n}\cos n\pi=\frac{-\sqrt{\pi}}{n}(-1)^n.$$

Hence from equation (I.2)

$$f(x)=\frac{\pi}{4}+\sum_{n=1}^{\infty}\left\{\frac{1}{\pi n^2}[(-1)^n-1]\cos nx-\frac{1}{n}(-1)^n\sin nx\right\}.$$

I.5. We have generally for m and n zero or positive integers, that

$$(x^m, x^n)=\int_{-1}^{1}x^{m+n}dx=\frac{1}{m+n+1}\Big[x^{m+n+1}\Big]_{-1}^{+1}=\begin{cases}\dfrac{2}{m+n+1}, & \text{if } (m+n)=\text{even}\\ 0, & \text{if } (m+n)=\text{odd}.\end{cases}$$

Hence $(\varphi_0, \varphi_0)=2$, *i.e.* $\psi_0=\frac{1}{\sqrt{2}}\varphi_0=\frac{1}{\sqrt{2}}$. We next let

$$\chi_1=\varphi_1-\psi_0(\psi_0, \varphi_1)=\varphi_1=x,$$

hence $(\chi_1, \chi_1)=\frac{2}{3}$, *i.e.* $\psi_1=\sqrt{\frac{3}{2}}\varphi_1=\sqrt{\frac{3}{2}}x$. We next let

$$\chi_2=\varphi_2-\psi_1(\psi_1,\varphi_2)-\psi_0(\psi_0,\varphi_2)$$
$$=\varphi_2-\frac{\sqrt{2}}{3}\psi_0=x^2-\frac{1}{3}$$

hence

$$(\chi_2,\chi_2)=(x^2,x^2)-\frac{2}{3}(x^2,1)+\frac{1}{9}(1,1)=\frac{8}{45}$$

and therefore

$$\psi_2=\sqrt{\frac{5}{2}}\left(\frac{3}{2}x^2-\frac{1}{2}\right).$$

By considering further functions $\varphi_3=x^3$, $\varphi_4=x^4$, . . . we can generate a set of normed orthogonal polynomials, ψ_0, ψ_1, ψ_2, . . . which form a complete set of functions for the interval $-1\leqslant x\leqslant 1$. These polynomials are called Legendre polynomials. They are of great importance in quantum mechanics.

I.6. From (I.6c)

$$\int_{-\infty}^{\infty}|f(x)|^2dx=\int_{-L}^{L}|e^{ik_0x}|^2dx=2L,$$

hence the Fourier integral theorem is applicable. (I.6b) then gives

$$c(k)=\frac{1}{2\pi}\int_{-L}^{L}e^{i(k_0-k)x}dx=\frac{\sin(k_0-k)L}{\pi(k_0-k)}. \qquad \text{(H 2)}$$

Equation (I.6a) can be interpreted as follows. e^{ikx} is a plane wave of wave-number $k/2\pi$, *i.e.* wavelength $\lambda=2\pi/k$. $f(x)$ represents a superposition of plane waves e^{ikx} with amplitudes $c(k)$. Given the resulting amplitude $f(x)$, equation (I.6b) gives the amplitude of any one monochromatic component e^{ikx}.

Equation (I.6c) represents a sinusoidal wave-train of length $2L$ and wavelength $2\pi/k_0$. In practice we have of course always only wave-trains of finite length, never infinitely long wave-trains. Equation (H 2), or better the intensity distribution

$$|c(k)|^2=\frac{\sin^2(k_0-k)L}{\pi^2(k_0-k)^2}, \qquad \text{(H 3)}$$

shows that such finite wave-trains are never strictly monochromatic. The intensity distribution (H 3) is indicated in *Figure 15.* The distribution has a maximum at $k=k_0$,

$$|c(k_0)|^2=\frac{L^2}{\pi^2}.$$

The first minima occur at $k=k_0\pm\frac{\pi}{L}$,

$$\left|c\left(k_0\pm\frac{\pi}{L}\right)\right|^2=0.$$

On account of the factor $(k_0 - k)^2$ in the denominator of (H 3), contributions to the amplitude $f(x)$ by wave-numbers k with $|k - k_0| > \pi/L$ are very small. Practically all the contributions come from the band of wave-numbers $k_0 - (\pi/L) < k < k_0 + (\pi/L)$, *i.e.* we have an intensity distribution of approximate spread $2\pi/L$*. Increasing L, the width

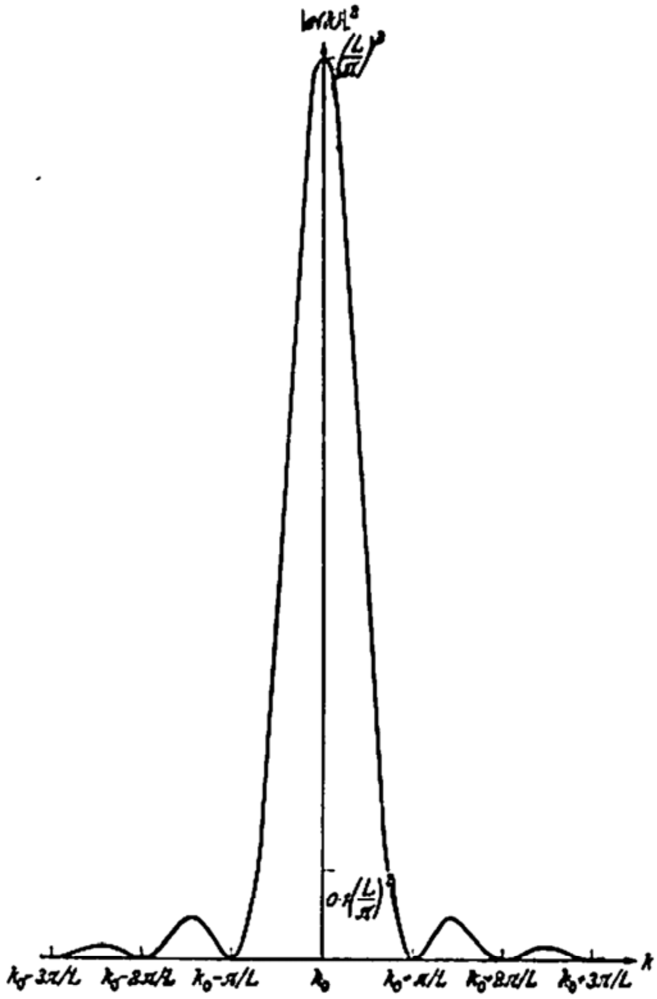

Figure 15. Spectral intensity distribution of sinusoidal wave-train of finite length $2L$ and wavelength $2\pi/k_0$

of this waveband decreases; for $L \gg \lambda_0 = 2\pi/k_0$, *i.e.* a wave-train many wavelengths long, we obtain a very sharp, narrow distribution. The above analysis and its interpretation is well illustrated in optics. Spectral lines due to finite wave-trains are not infinitely sharp, but have a finite, if small, width.

---

* In fact, writing $x = (k - k_0)L$, one finds by numerical integration that

$$\int_{k_0-\pi/L}^{k_0+\pi/L} |c(k)|^2 dk = \frac{L}{\pi}\int_{-\pi}^{\pi} \frac{\sin^2 x}{\pi x^2} dx = 0{\cdot}90\,\frac{L}{\pi},$$

whereas

$$\int_{-\infty}^{\infty} |c(k)|^2 dk = \frac{L}{\pi}\int_{-}^{\infty} \frac{\sin^2 x}{\pi x^2} dx = \frac{L}{\pi}.$$

I.7. From $f=(\varphi,\psi)\varphi-(\varphi,\varphi)\psi$, it follows that $(f,f)\geqslant 0$ becomes

$$\begin{aligned}(f,f)&=((\varphi,\psi)\varphi-(\varphi,\varphi)\psi,\ (\varphi,\psi)\varphi-(\varphi,\varphi)\psi)\\&=\overline{(\varphi,\psi)}(\varphi,\psi)(\varphi,\varphi)-\overline{(\varphi,\psi)}(\varphi,\varphi)(\varphi,\psi)\\&\quad-(\varphi,\varphi)(\varphi,\psi)(\psi,\varphi)+(\varphi,\varphi)(\varphi,\varphi)(\psi,\psi)\\&=(\varphi,\varphi)\{(\varphi,\varphi)(\psi,\psi)-(\varphi,\psi)\overline{(\varphi,\psi)}\}\\&=(\varphi,\varphi)\{\|\varphi\|^2\|\psi\|^2-|(\varphi,\psi)|^2\}\geqslant 0\end{aligned}$$

and since $(\varphi,\varphi)>0$, Schwarz's inequality follows:

$$|(\varphi,\psi)|\leqslant\|\varphi\|\ \|\psi\|.$$

The equality sign holds, *i.e.* $(f,f)=0$, if and only if

$$f=(\varphi,\psi)\varphi-(\varphi,\varphi)\psi=0$$

i.e.

$$\psi=\frac{(\varphi,\psi)}{(\varphi,\varphi)}\varphi=a\varphi,$$

where a is a complex number.

From equation (1.4*) one has at once, that for an ordinary vector space,

$$|(\mathbf{u},\mathbf{v})|=\|\mathbf{u}\|\ \|\mathbf{v}\|\ |\cos\theta|\leqslant\|\mathbf{u}\|\ \|\mathbf{v}\|.$$

Schwarz's inequality is the generalization of this result to Hilbert space.

I.8. The equation of a vibrating membrane is

$$\frac{1}{c^2}\frac{\partial^2u}{\partial t^2}=\frac{\partial^2u}{\partial x^2}+\frac{\partial^2u}{\partial y^2}.$$

We obtain equation (I.8a) and the boundary conditions (I.8b) by writing

$$u(x,y,t)=w(x,y)T(t)$$

and separating the variables. In (I.8a) we again separate the variables,

$$w(x,y)=X(x)Y(y),$$

and obtain two eigenvalue problems for $X(x)$ and $Y(y)$ identical with (2.5). We then get for the eigensolutions and corresponding eigenvalues

$$w_{m,n}(x,y)=\sin\frac{m\pi}{L}x\sin\frac{n\pi}{L}y$$

$$\lambda_{m,n}=\left(\frac{\pi}{L}\right)^2(m^2+n^2),$$

where m and n are restricted to the values

$$m,n=1,2,3,\ldots$$

The eigenvalues $\lambda_{m,n}$ are two-fold degenerate for $m \neq n$ corresponding to the two different solutions $w_{n,m}$ and $w_{m,n}$. Hence any linear combination

$$aw_{m,n} + bw_{n,m}$$

is also an eigenfunction of the problem belonging to the same eigenvalues $\lambda_{n,m} = \lambda_{m,n}$.

CHAPTER II

II.1. The classical angular momentum vector $\mathbf{L}$ is

$$\mathbf{L} = \sum_{s=1}^{n} \mathbf{r}_s \wedge \mathbf{p}_s.$$

Replacing the classical quantities by the corresponding operators [equations (6.1) and (6.2)]

$$\mathbf{L}_{op} = \sum_{s=1} \mathbf{r}_s \wedge (-i\hbar \operatorname{grad}_s)$$

or for the components

$$(L_x)_{op} = -i\hbar \sum_{s=1}^{n} \left(y_s \frac{\partial}{\partial z_s} - z_s \frac{\partial}{\partial y_s} \right), \ etc.$$

II.2. The classical conservation of energy equation is

$$\frac{1}{2m}(\mathbf{p}_1{}^2 + \mathbf{p}_2{}^2) + \frac{1}{2M}\mathbf{P}^2 - 2e^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{e^2}{r_{12}} = E,$$

where $\mathbf{r}_1$, $\mathbf{r}_2$ are the position vectors of the two electrons relative to the nucleus, $r_1 = |\mathbf{r}_1|$, *etc*, $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$; m, M are the masses of electron and nucleus respectively, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{P}$ are the momenta of the electrons and nucleus respectively.

Using equations (6.1)–(6.3) we obtain the time-dependent Schrödinger equation for $u(\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}, t)$, $\mathbf{R}$ being the position coordinate of the nucleus,

$$\left[-\frac{\hbar^2}{2m}(\nabla_1{}^2 + \nabla_2{}^2) - \frac{\hbar^2}{2M}\nabla^2 - 2e^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{e^2}{r_{12}} \right] u(\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}, t) = i\hbar \frac{\partial}{\partial t} u(\mathbf{r}_1, \mathbf{r}_2, \mathbf{R}, t),$$

where $\nabla_1{}^2 = \partial^2/\partial x_1{}^2 + \partial^2/\partial y_1{}^2 + \partial^2/\partial z_1{}^2$ *etc*,

and $\nabla^2 = \partial^2/\partial X^2 + \partial^2/\partial Y^2 + \partial^2/\partial Z^2$.

If the nucleus were infinitely heavy ($M/m \to \infty$) the kinetic energy term ($-\hbar^2\nabla^2/2M$) would not occur in the Schrödinger equation. Since $M/m \simeq 7400$, one can to very good approximation neglect this term.

II.3. With $E=E(p)$, $E_0=E(p_0)$, we put

$$px-Et=(p_0x-E_0t)+\{(p-p_0)x-(E-E_0)t\}$$

and can then write equation (II.3) as

$$u(x,t)=A\exp\frac{i}{\hbar}(p_0x-E_0t),$$

where

$$A=\int_{p_0-\Delta p_0/2}^{p_0+\Delta p_0/2}a(p)\exp\left[\frac{i}{\hbar}\{(p-p_0)x-(E-E_0)t\}\right]dp.$$

By analogy with a plane wave we call A the amplitude of the wave packet. Surfaces of constant amplitude are defined by

$$(p-p_0)x-(E-E_0)t=\text{const.},$$

and they are propagated with the velocity

$$c_g=\frac{dx}{dt}=\frac{E-E_0}{p-p_0}.$$

For a sufficiently sharp wave-packet, $(E-E_0)/(p-p_0)$ approximates the differential coefficient at $p=p_0$ arbitrarily closely,

$$c_g=\left(\frac{dE}{dp}\right)_{p=p_0}.$$

Using equations (5.1) and (5.3), equation (6.8) follows.

II.4. Equation (II.4) follows at once from equation (6.13), by applying Gauss's theorem to the surface integral, and using the fact that (6.13) holds for *any* domain of integration V.

II.5. For a wave-train of length $2L$, equation (I.6c), the uncertainty in position of the particle is

$$\Delta x\simeq 2L. \tag{H 4}$$

From the discussion following equation (H 3), and *Figure 15*, the uncertainty in the wave-number k is of the order of magnitude of $2\pi/L$, *i.e.* since $p=\hbar k$,

$$\Delta p=\hbar\Delta k\simeq\hbar 2\pi/L=h/L \tag{H 5}$$

and from (H 4) and (H 5)

$$\Delta x\,.\,\Delta p\simeq 2h.$$

CHAPTER III

III.1. We shall write $(\mathbf{r},\mathbf{p})$ for $(\mathbf{r}_1,\ldots\mathbf{r}_n,\mathbf{p}_1,\ldots\mathbf{p}_n)$. $\psi(\mathbf{r})$ satisfies the Schrödinger equation

$$H(\mathbf{r},\mathbf{p})\psi(\mathbf{r})=E\psi(\mathbf{r}). \tag{H 6}$$

Under inversion, $\mathbf{r} \to -\mathbf{r}$, and (remembering that $\mathbf{p} = -i\hbar\,\mathrm{grad}$, so that $\mathbf{p} \to -\mathbf{p}$ under inversion), equation (H 6) transforms into

$$H(-\mathbf{r}, -\mathbf{p})\psi(-\mathbf{r}) = E\psi(-\mathbf{r}).$$

On account of equation (III.1), this equation is the same as

$$H(\mathbf{r}, \mathbf{p})\psi(-\mathbf{r}) = E\psi(-\mathbf{r}) \tag{H 7}$$

i.e. $\psi(-\mathbf{r})$ is also an eigenfunction of H belonging to the same energy eigenvalue E. But since we assume that E is non-degenerate, $\psi(-\mathbf{r})$ must be essentially the same function as $\psi(\mathbf{r})$, *i.e.* it must be a multiple of $\psi(\mathbf{r})$:

$$\psi(-\mathbf{r}) = \lambda\psi(\mathbf{r}), \tag{H 8}$$

applying (H 8) twice gives

$$\psi(\mathbf{r}) = \lambda^2\psi(\mathbf{r}),$$

i.e. $\lambda^2 = 1$, *i.e.*

$$\lambda = \pm 1.$$

If we allow E to be a degenerate eigenvalue then we cannot deduce equation (H 8), but equation (H 7) still holds: $\psi(\mathbf{r})$ and $\psi(-\mathbf{r})$, and consequently any linear combination of them, are eigenfunctions of H belonging to the same eigenvalue E. We can then choose two linear combinations, which span the same manifold as $\psi(\mathbf{r})$ and $\psi(-\mathbf{r})$, and which have a definite parity. These two linear combinations are

$$\psi(\mathbf{r}) \pm \psi(-\mathbf{r}).$$

III.2. The probability current vector $\mathbf{j}$ is given by equation (6.12):

$$\mathbf{j} = \frac{\hbar}{2mi}(\bar{\psi}\,\mathrm{grad}\,\psi - \psi\,\mathrm{grad}\,\bar{\psi}).$$

For $\psi = e^{ikr}/r$, the only non-vanishing component of $\mathbf{j}$ is the radial component

$$j_r = \frac{\hbar k}{m}\frac{1}{r^2} = \frac{v}{r^2}$$

where v is the magnitude of the particle velocity. Thus we are dealing with an isotropic wave of outgoing particles; the number of particles crossing a sphere $r = \text{const.}$, per unit time, is $4\pi v$.

If we had considered $\psi = e^{-ikr}/r$ we would have found

$$j_r = -\frac{v}{r^2},$$

i.e. an isotropic wave of incoming particles.

III.3. Let $M = m_1 + m_2$, then

$$\mathbf{R} = \frac{m_1}{M}\mathbf{r}_1 + \frac{m_2}{M}\mathbf{r}_2, \quad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2.$$

From these equations one easily proves that

$$\frac{1}{m_1}\nabla_1^2 + \frac{1}{m_2}\nabla_2^2 = \frac{1}{M}\nabla_R^2 + \frac{1}{\mu}\nabla^2,$$

where $\mathbf{R}$ and $\mathbf{r}$ have components (X, Y, Z) and (x, y, z), and

$$\nabla_R^2 = \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}, \quad \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

and μ is the *reduced mass* of the two masses m_1 and m_2, defined by

$$\mu = \frac{m_1 m_2}{m_1 + m_2}.$$

Writing $\psi(\mathbf{r}_1, \mathbf{r}_2, t) = \Phi(\mathbf{R}, t)\varphi(\mathbf{r}, t)$, the Schrödinger equation (III.3) is separable, giving

$$i\hbar\frac{\partial}{\partial t}\Phi(\mathbf{R}, t) = -\frac{\hbar^2}{2M}\nabla_R^2\Phi(\mathbf{R}, t) \qquad \text{(H 9)}$$

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r}, t) = -\frac{\hbar^2}{2\mu}\nabla^2\varphi(\mathbf{r}, t) + V(\mathbf{r})\varphi(\mathbf{r}, t). \qquad \text{(H 10)}$$

Equation (H 9) is the Schrödinger equation of a free particle of mass M. It represents the motion of the centre of gravity relative to the (arbitrary) coordinate system. Equation (H 10) describes the relative motion of the two particles. It represents the motion of a particle of mass μ in the potential field $V(\mathbf{r})$. This separation into centre of mass and relative motion is the same as occurs in classical mechanics.

III.4. The problem is treated in essentially the same manner as the problem in section 8. The constants of integration in the three regions $x<0$, $0<x<a$, $a<x$, must be determined by matching the wave functions and their first derivatives at $x=0$ and at $x=a$.

For the case $E>V_0$, this leads to

$$R = \left[1 + \frac{4E(E-V_0)}{V_0^2 \sin^2(k_2 a)}\right]^{-1}, \quad T = \left[1 + \frac{V_0^2 \sin^2(k_2 a)}{4E(E-V_0)}\right]^{-1},$$

where k_2 is defined in equation (8.3). This problem has been considered in detail by Schiff[12] (section 17).

III.5. The solutions must be of the form

$$\psi = Ae^{\kappa x}, \quad \kappa = \left\{\frac{2m}{\hbar^2}(V_0 - E)\right\}^{\frac{1}{2}}, \quad \text{when } x < -a,$$

$$\psi = Be^{ikx} + Ce^{-ikx}, \quad k = \left\{\frac{2m}{\hbar^2}E\right\}^{\frac{1}{2}} \quad \text{when } |x| < a,$$

$$\psi = De^{-\kappa x}, \qquad \text{when } x > a.$$

These solutions and their first derivatives must be equated at $x = \pm a$ to determine the constants of integration, apart from a common factor determined from the normalization of the wave function.

(*i*) Using the fact that the eigensolutions must be even or odd, one can straight away say that there are two types of solutions; both behave like

$$\psi = De^{-\kappa x} \quad \text{when } x > a.$$

For $0<x<a$, they behave like

$$\psi = B \cos kx \quad \text{and} \quad \psi = C \sin kx$$

respectively.

(*ii*) Matching logarithmic derivatives for the even eigenfunctions gives at once

$$k \tan ka = \kappa.$$

This is a transcendental equation in which all quantities except E are known. Its solutions give the allowed discrete set of energy eigenvalues.

For the odd eigenfunctions the situation is similar, the equation for the energy eigenvalues being

$$k \cot ka = -\kappa.$$

This problem has been discussed in detail in Schiff[12] (pp. 36–40); in particular, a simple method for solving the above transcendental equations is given.

III.6. The wave equation of the hydrogen atom is given by (12.1) with $V(r) = -e^2/r$. For m one should strictly take the reduced mass

$$m = \frac{m_{\text{electron}} \cdot m_{\text{proton}}}{m_{\text{electron}} + m_{\text{proton}}},$$

but since $m_{\text{proton}}/m_{\text{electron}} \simeq 1840$, $m \simeq m_{\text{electron}}$ in all atomic problems.

Spherically symmetric solutions, *i.e.* solutions depending on r only, are obtained from (12.7) with $\lambda = 0$. For the ground state, the electron is bound, *i.e.* $E<0$. Hence we put $\alpha^2 = -2mE/\hbar^2$. The approximate asymptotic solution (12.10) now becomes $e^{-\alpha r}$. Substituting this in (12.7), one finds that it represents an exact solution, provided that

$$\alpha = \frac{me^2}{\hbar^2};$$

the normed wave-function, derived from

$$\int_{\substack{\text{all}\\\text{space}}} |\psi_0|^2 d\tau = 1,$$

is

$$\psi_0 = \left(\frac{\alpha^3}{\pi}\right)^{\frac{1}{2}} e^{-\alpha r}.$$

It represents the ground state since it contains no nodes. The corresponding energy level, *i.e.* the ionization energy of the hydrogen atom is

$$E_0 = -\frac{me^4}{2\hbar^2} = -13{\cdot}5 \text{ eV}$$

(the electron volt (eV) being the energy acquired by an electron in falling through a potential difference of 1 volt. 1 eV $= 1{\cdot}59 \times 10^{-12}$ erg). $1/\alpha$ is called the Bohr radius of the hydrogen atom. It is the most probable distance between proton and electron as one can easily verify by finding the maximum of $r^2|\psi_0|^2$. $1/a = 0{\cdot}528 \times 10^{-8}$ cm, 10^{-8} cm being typical for the order of magnitude of atomic dimensions.

The corresponding wave function of an electron moving in the Coulomb field of a nucleus of charge Ze, neglecting the effect of any other electrons present, is obtained by replacing e^2 by Ze^2 everywhere, *i.e.*

$$\psi_0 = \left(\frac{Z^3 a^3}{\pi}\right)^{\frac{1}{2}} e^{-Zar}.$$

This is, for example, the correct wave function for singly-ionized helium ($Z = 2$), or doubly-ionized lithium ($Z = 3$).

If we neglect the Coulomb repulsion between the two electrons in helium, the wave equation for helium is separable and the ground state wave function is simply a product of two wave functions of the above kind with $Z = 2$:

$$\psi_0(\mathbf{r}_1, \mathbf{r}_2) = \left(\frac{8a^3}{\pi}\right) e^{-2a(r_1 + r_2)}.$$

CHAPTER IV

IV.1. From equations (IV.1) and (15.5)

$$\dot{q}_i = \frac{1}{m} p_i, \quad \dot{p}_i = -\frac{\partial V}{\partial q_i}, \quad (i = 1, 2, 3).$$

Equation (IV.1) is the Hamiltonian function for a particle of mass m, with cartesian coordinates (q_1, q_2, q_3), moving in a potential field $V(q_1, q_2, q_3)$. Hamilton's equations merely state the conventional definition of momentum and Newton's law of motion.

IV.2. From (12.2), (15.1) and (15.2),

$$T = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta\, \dot{\varphi}^2)$$
$$V = V(r).$$

Hence from (15.3)

$$\left.\begin{aligned} p_r &= \frac{\partial T}{\partial \dot{r}} = m\dot{r} \\ p_\theta &= \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta} \\ p_\varphi &= \frac{\partial T}{\partial \dot{\varphi}} = mr^2 \sin^2 \theta\, \dot{\varphi}. \end{aligned}\right\}$$

The Hamiltonian function is

$$H = \frac{1}{2m}\left[p_r^2 + \frac{1}{r^2} p_\theta^2 + \frac{1}{r^2 \sin^2 \theta} p_\varphi^2\right] + V(r)$$

and Hamilton's equations follow at once by elementary differentiation

$$\dot{r} = \frac{1}{m} p_r$$

$$\dot{p}_r = -\frac{\partial V}{\partial r} + \frac{1}{mr^3}\left[p_\theta{}^2 + \frac{1}{\sin^2\theta} p_\varphi{}^2\right]$$

etc.

[Note that the generalized momenta do not necessarily have the dimensions of momentum, *i.e.* of

(mass × length)/time;

from (15.3) it follows that the dimensions of p_i are such that $p_i \dot{q}_i$ is an energy. Thus p_r has the dimensions of momentum, p_θ and p_φ those of angular momentum.]

IV.3. Let the operator A have the complete normed orthogonal sequence of eigenfunctions $\varphi_1, \varphi_2, \ldots, \varphi_n, \ldots$ and let $a_1 \leqslant a_2 \leqslant \ldots \leqslant a_n \leqslant \ldots$ be the corresponding real eigenvalues. If ψ and χ are any two vectors of Hilbert space, they can be expanded in terms of the φ's

$$\psi = \sum_{n=1}^{\infty} c_n \varphi_n, \quad \chi = \sum_{n=1}^{\infty} d_n \varphi_n;$$

hence

$$(\chi, A\psi) = \left(\sum_{m=1}^{\infty} d_m \varphi_m, \sum_{n=1}^{\infty} a_n c_n \varphi_n\right) = \sum_{n=1}^{\infty} a_n c_n \bar{d}_n$$

and, since $\bar{a}_n = a_n$,

$$(A\chi, \psi) = \left(\sum_{m=1}^{\infty} a_m d_m \varphi_m, \sum_{n=1}^{\infty} c_n \varphi_n\right) = \sum_{n=1}^{\infty} a_n c_n \bar{d}_n$$

i.e. $(A\chi, \psi) = (\chi, A\psi)$, *i.e.* A is Hermitian.

IV.4. From equation (16.19)

$$(\Delta l_z)^2 = \langle l_z{}^2 \rangle,$$

since $\langle l_z \rangle = 0$. And from equations (16.20), (14.5), (13.21) and

$$\int_0^\infty |R(r)|^2 r^2 dr = 1$$

it follows that

$$\langle l_z{}^2 \rangle = (l_z\psi, l_z\psi) = \frac{1}{2l+1}\hbar^2 \sum_{m=-l}^{l} m^2 = \frac{2\hbar^2}{(2l+1)} \sum_{m=1}^{l} m^2$$

and since $\sum_{m=1}^{l} m^2 = \frac{1}{6} l(l+1)(2l+1)$,

$$\langle l_z{}^2 \rangle = \tfrac{1}{3}\hbar^2 l(l+1).$$

IV.5. (*i*) Since $(U^\star)^\star = U$, equations (18.14) are also the conditions for $U^\star$ to be unitary: they are symmetrical in U and $U^\star$;

(*ii*) $(UV)(UV)^\star = UVV^\star U^\star = UU^\star = \mathbf{I}$, *etc*;
and similarly, since $(V^{-1})^\star = V$,

(*iii*) $(UV^{-1})(UV^{-1})^\star = UV^{-1}VU^\star = \mathbf{I}$, *etc*.

IV.6. If X is any inverse operator of A, then $XA = \mathbf{I}$ and $AX = \mathbf{I}$. Multiplying the first of these equations by A^{-1} on the right-hand side and using (18.18), $X = A^{-1}$ follows, *i.e.* A^{-1} is unique.

Similarly, if $X = (AB)^{-1}$ is the inverse operator of AB, then $ABX = \mathbf{I}$ and $XAB = \mathbf{I}$. Multiplying the first of these equations successively by A^{-1} and B^{-1} on the left-hand side one obtains $BX = A^{-1}$, and $X = B^{-1}A^{-1}$, and this X obviously also satisfies the second equation $XAB = \mathbf{I}$. Hence

$$(AB)^{-1} = B^{-1}A^{-1}.$$

This relation was to be expected, since a similar relation, equation (18.9), holds for the adjoint operators, and for unitary operators, for which $U^\star = U^{-1}$, the two equations must agree.

IV.7. From equations (18.16) and (18.17) it follows that the vectors $\psi_1, \psi_2, \dots \psi_n, \dots$ are normalized and orthogonal. They are also a complete set. For if they were not complete there would exist a vector f orthogonal to every member of the set, $(f, \psi_n) = 0$, for all n; *i.e.* $(U^{-1}f, \varphi_n) = 0$ for all n; in other words, there would exist a vector, $U^{-1}f$, orthogonal to all φ_n, in contradiction to the assumption that the φ_n are complete.

IV.8. $[p_r, p_s] = 0$, since for any wave function ψ, $\dfrac{\partial^2}{\partial x_r \partial x_s}\psi = \dfrac{\partial^2}{\partial x_s \partial x_r}\psi$.
Equation (5b) follows from

$$-i\hbar x_r \frac{\partial}{\partial x_s}\psi + i\hbar \frac{\partial}{\partial x_s}(x_r\psi) = i\hbar\psi \frac{\partial x_r}{\partial x_s} = i\hbar \delta_{rs}\psi.$$

IV.9. From equations (14.2) and (14.3), the required result follows at once. For the operator $\partial/\partial\varphi$ commutes with both position variables r and θ, but not with φ. But φ does not appear in (14.3).

IV.10. From Exercise III.1 we know that $\psi(x)$ has a definite parity, so that $|\psi(x)|^2 = |\psi(-x)|^2$. Now

$$\langle x \rangle = \int_{-\infty}^{\infty} x|\psi(x)|^2 dx$$

and the integrand is an odd function of x whereas the range of integration is $-\infty < x < \infty$, so that $\langle x \rangle = 0$.

IV.11. To find the momentum distribution we represent $\psi(\mathbf{r})$ as a superposition of eigenfunctions of definite momentum, *i.e.* of plane waves. The correctly normalized plane waves are

$$\varphi(\mathbf{p}, \mathbf{r}) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}} e^{i\mathbf{p}\mathbf{r}/\hbar} \tag{H 11}$$

so that, according to (4.17) and (4.18),

$$\int \overline{\varphi(\mathbf{p}', \mathbf{r})}\varphi(\mathbf{p}, \mathbf{r})d\mathbf{r} = \delta(\mathbf{p}' - \mathbf{p}), \qquad \text{(H 12)}$$

the triple integral being over all values of $\mathbf{r}$, *i.e.* $-\infty < x < \infty$, *etc.* We then write

$$\psi(\mathbf{r}) = \int A(\mathbf{p})\varphi(\mathbf{p}, \mathbf{r})d\mathbf{p}, \qquad \text{(H 13)}$$

the triple integral being over all values of $\mathbf{p}$, *i.e.* $-\infty < p_x < \infty$, *etc.* According to (17.2) the probability of a particle having a momentum in the range $(\mathbf{p}, \mathbf{p}+d\mathbf{p})$ is $|A(\mathbf{p})|^2 d\mathbf{p}$. From (H 11), (H 12) and (H 13) one easily finds $A(\mathbf{p})$ explicitly,

$$A(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{\frac{3}{2}}}\int_{\substack{\text{all}\\ \text{space}}} \psi(\mathbf{r})e^{-i\mathbf{p}\mathbf{r}/\hbar}d\mathbf{r}. \qquad \text{(H 14)}$$

This equation could have been derived directly, without introducing the δ-function, from the Fourier integral theorem (*cf.* Exercise I.6); in fact, the momentum distribution is the Fourier transform of the wave function. Using the above equations, we can easily check that we used correctly normed momentum wave functions, *i.e.* that the momentum distribution is correctly normed, *i.e.*

$$\int |A(\mathbf{p})|^2 d\mathbf{p} = 1.$$

For

$$\begin{aligned}
\int |A(\mathbf{p})|^2 d\mathbf{p} &= \frac{1}{(2\pi\hbar)^3}\iiint \overline{\psi(\mathbf{r})}\psi(\mathbf{s})e^{i\mathbf{p}(\mathbf{r}-\mathbf{s})/\hbar}d\mathbf{r}d\mathbf{s}d\mathbf{p} \\
&= \int \overline{\psi(\mathbf{r})}\psi(\mathbf{s})\delta(\mathbf{r}-\mathbf{s})d\mathbf{r}d\mathbf{s} \\
&= \int |\psi(\mathbf{r})|^2 d\mathbf{r} = 1
\end{aligned}$$

since $\psi(\mathbf{r})$ was assumed normalized.

IV.12.

$$\begin{aligned}
[A, B+C] &= A(B+C) - (B+C)A = [A, B] + [A, C] \\
[A, BC] &= ABC - BCA \\
&= (AB - BA)C + B(AC - CA) \\
&= [A, B]C + B[A, C].
\end{aligned}$$

The third relation (IV.12c) follows from (IV.12b) by inverting the order of the factors, since $[A, B] = -[B, A]$.

IV.13. Using the commutation relations (19.5) and Exercise IV.12, equations (IV.13a–d) are easily proved. Some typical proofs are:

$$\begin{aligned}
[l_x, x] &= [yp_z - zp_y, x] = 0\\
[l_x, y] &= [yp_z - zp_y, y] = -z[p_y, y] = i\hbar z\\
[l_x, p_x] &= [yp_z - zp_y, p_x] = 0\\
[l_x, p_y] &= [yp_z - zp_y, p_y] = [y, p_y]p_z = i\hbar p_z\\
[l_x, l_y] &= [yp_z - zp_y, l_y] = y[p_z, l_y] - [z, l_y]p_y\\
&= y(-i\hbar p_x) - (-i\hbar x)p_y\\
&= i\hbar(xp_y - yp_x)\\
&= i\hbar l_z.\\
[l_x, \mathbf{l}^2] &= [l_x, l_y^2] + [l_x, l_z^2]\\
&= [l_x, l_y]l_y + l_y[l_x, l_y] + [l_x, l_z]l_z + l_z[l_x, l_z]\\
&= i\hbar l_z l_y + l_y i\hbar l_z + (-i\hbar l_y)l_z + l_z(-i\hbar l_y) = 0.
\end{aligned}$$

CHAPTER V

V.1. If $[A_{mn}] = [a_n\delta_{mn}]$, $B_{mn} = [b_n\delta_{mn}]$, then

$$(AB)_{mn} = \sum_{k=1}^{N} A_{mk}B_{kn} = \sum_{k=1}^{N} a_m b_k \delta_{mk}\delta_{kn} = a_n b_n \delta_{mn};$$

this is a diagonal matrix and is symmetric in A and B, hence $AB = BA$.

V.2 From (22.19) one gets

$$\sum_{m,n=1}^{N} A_{mn}c_m\bar{c}_n = a\sum_{n=1}^{N}|c_n|^2,$$

and taking complex conjugates of this equation, since $A_{mn} = \bar{A}_{nm}$:

$$\bar{a}\sum_{n=1}^{N}|c_n|^2 = \sum_{m,n=1}^{N}\bar{A}_{mn}\bar{c}_m c_n = \sum_{m,n=1}^{N}A_{nm}\bar{c}_m c_n = \sum_{m,n=1}^{N}A_{mn}\bar{c}_n c_m,$$

interchanging the summation suffixes m and n. Hence $a = \bar{a}$.

V.3. Suppose

$$\sum_{m=1}^{N}A_{nm}c_m = a_1 c_n, \quad \sum_{m=1}^{N}A_{nm}d_m = a_2 d_n.$$

Hence

$$\sum_{m,n=1}^{N}A_{nm}c_m\bar{d}_n = a_1\sum_{n=1}^{N}c_n\bar{d}_n, \tag{H 15}$$

and

$$\sum_{m,n=1}^{N}\bar{A}_{nm}\bar{d}_m c_n = a_2\sum_{n=1}^{N}\bar{d}_n c_n, \quad (\text{since } \bar{a}_2 = a_2),$$

i.e.

$$\sum_{m,n=1}^{N}A_{nm}c_m\bar{d}_n = a_2\sum_{n=1}^{N}\bar{d}_n c_n \tag{H 16}$$

and subtracting equations (H 15) and (H 16)

$$0=(a_1-a_2)\sum_{n=1}^{N}\bar{d}_n c_n,$$

hence if $a_1 \neq a_2$, $[c_n]$ is orthogonal to $[d_n]$.

V.4. From the law of matrix multiplication (22.2)

$$s_x{}^2=\frac{\hbar^2}{4}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=\frac{\hbar^2}{4}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\tfrac{1}{4}\hbar^2\,\mathbf{I}$$

and similarly $s_y{}^2=s_z{}^2=\frac{1}{4}\hbar^2\mathbf{I}$; hence $\mathbf{s}^2=\frac{3}{4}\hbar^2\mathbf{I}$. Thus the operators s_z and $\mathbf{s}^2$ are diagonal and hence simultaneously observable. s_z has the eigenvalues $\pm\hbar/2$, $\mathbf{s}^2$ has the degenerate eigenvalue $\frac{3}{4}\hbar^2$.

One easily verifies that the operators s_x, s_y, s_z do not commute, *e.g.*

$$s_x s_y=\frac{\hbar^2}{4}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}=\frac{\hbar^2 i}{4}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

$$s_y s_x=\frac{\hbar^2}{4}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=\frac{\hbar^2 i}{4}\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix};$$

hence

$$[s_x, s_y]=i\hbar s_z,$$

and two similar relations obtained by cyclic interchange of x, y and z. These commutation relations, like the above properties of these operators which we have already met, are generally true for angular momentum operators not only for the particular 2×2 representation we used above.

V.5. Let $U_{mn}(m, n=1, 2)$ be the 2×2 matrix which transforms the vectors ξ_1, ξ_2 into the eigenfunctions φ_1, φ_2 of A, *i.e.* from equation (22.23),

$$\varphi_n=U_{1n}\xi_1+U_{2n}\xi_2, \quad (n=1, 2).$$

The eigenvalue problem (22.21) is in our case

$$(\lambda-a_n)U_{1n}+\mu U_{2n}=0$$
$$\mu U_{1n}+(\lambda-a_n)U_{2n}=0.$$

Each of these equations gives the ratio U_{1n}/U_{2n}. They are consistent if these two ratios are the same, *i.e.* if

$$\frac{\lambda-a_n}{\mu}=\frac{\mu}{\lambda-a_n}. \tag{H 17}$$

[This equation is just the consistency condition (22.22),

$$\begin{vmatrix}\lambda-a_n & \mu\\ \mu & \lambda-a_n\end{vmatrix}=0.]$$

From the quadratic equation in a_n, (H 17), we obtain two solutions

$$a_1=\lambda-\mu, \quad a_2=\lambda+\mu$$

as the eigenvalues of A. Corresponding to a_1 we find $U_{11}/U_{21} = -1$, and corresponding to a_2 we find $U_{12}/U_{22} = +1$. For U to be unitary we must have $|U_{11}|^2 + |U_{12}|^2 = 1$ and $|U_{21}|^2 + |U_{22}|^2 = 1$, so that we can take $U_{11} = U_{12} = 1/\sqrt{2}$. Hence the orthonormal eigensolutions are

$$\varphi_1 = \frac{1}{\sqrt{2}}(\xi_1 - \xi_2), \quad \varphi_2 = \frac{1}{\sqrt{2}}(\xi_1 + \xi_2).$$

V.6. One can prove (23.4) by the same method which is used in section 23. In the Schrödinger picture, the wave functions ψ_t and χ_t are time-dependent and

$$\frac{dA(t)}{dt} = \frac{\partial A(t)}{\partial t},$$

so that

$$\frac{d}{dt}(\chi_t, A\psi_t) = \left(\frac{\partial \chi_t}{\partial t}, A\psi_t\right) + \left(\chi_t, \frac{\partial A}{\partial t}\psi_t\right) + \left(\chi_t, A\frac{\partial \psi_t}{\partial t}\right)$$
$$= \left(\chi_t, \frac{\partial A}{\partial t}\psi_t\right) + \left(\chi_t, \frac{AH}{i\hbar}\psi_t\right) - \left(\chi_t, \frac{HA}{i\hbar}\psi_t\right),$$

so that

$$i\hbar\frac{d}{dt}(\chi_t, A\psi_t) = \left(\chi_t, \left\{i\hbar\frac{\partial A}{\partial t} + [A, H]\right\}\psi_t\right). \tag{H 18}$$

In the Heisenberg picture, the wave functions are always time-independent, ($\psi_t = \psi_0$, $\partial\psi_t/\partial t = 0$, *etc*) so that the whole time variation must be taken into account through $A(t)$:

$$i\hbar\frac{d}{dt}(\chi_t, A\psi_t) = i\hbar\left(\chi_t, \frac{dA}{dt}\psi_t\right). \tag{H 19}$$

Comparing equations (H 18) and (H 19) we see that the two modes of description are consistent if the Heisenberg equation of motion of the operator $A(t)$ is

$$i\hbar\frac{dA}{dt} = i\hbar\frac{\partial A}{\partial t} + [A, H].$$

V.7. The required matrix element is

$$\langle n|x|m\rangle = \int_{-\infty}^{\infty} \overline{\psi_n(x)}\, x\psi_m(x)\,dx = \frac{1}{\alpha^2}\int_{-\infty}^{\infty} \overline{\varphi_n(\xi)}\,\xi\varphi_m(\xi)\,d\xi$$
$$= \frac{1}{\alpha(2^{n+m}n!\,m!\,\pi)^{\frac{1}{2}}}\int_{-\infty}^{\infty} \mathrm{e}^{-\xi^2}\xi H_n(\xi)H_m(\xi)\,d\xi$$
$$= \frac{1}{\alpha(2^{n+m}n!\,m!\,\pi)^{\frac{1}{2}}}\int_{-\infty}^{\infty} \mathrm{e}^{-\xi^2}H_m(\xi)\,\{\tfrac{1}{2}H_{n+1}(\xi) + nH_{n-1}(\xi)\}\,d\xi \tag{H 20}$$

where we used (11.23) and the recurrence relation (11.19). From the orthogonality relation (11.22) it follows that the only non-vanishing terms in (H 20) occur for $m=n+1$ or $m=n-1$. Using (11.23) to obtain the correct normalization, one easily obtains

$$\langle n|x|m\rangle=\frac{1}{a}\left(\frac{n+1}{2}\right)^{\frac{1}{2}},\qquad (m=n+1),$$

$$\langle n|x|m\rangle=\frac{1}{a}\left(\frac{n}{2}\right)^{\frac{1}{2}},\qquad (m=n-1),$$

$$\langle n|x|m\rangle=0,\qquad (n-1\neq m\neq n+1).$$

CHAPTER VI

VI.1.

$$[A_+, M_z]=[M_x, M_z]+i[M_y, M_z]=-i\hbar M_y-\hbar M_x=-\hbar A_+,$$

$$[A_+, M^2]=0,\quad \text{from (25.7)},$$

$$\begin{aligned}A_+A_-=(M_x+iM_y)(M_x-iM_y)&=M_x{}^2+M_y{}^2-i(M_xM_y-M_yM_x)\\&=M^2-M_z{}^2+\hbar M_z.\end{aligned}$$

VI.2. From (25.23), if $\Psi(q, n, m)$ and $\Psi(q, n, m\pm 1)$ are to be normed:

$$\begin{aligned}|\mathcal{N}_\pm(q, n, m)|^2&=\langle q, n, m|A_\mp A_\pm|q, n, m\rangle\\&=\langle q, n, m|M^2-M_z{}^2\mp\hbar M_z|q, n, m\rangle\\&=\hbar^2n(n+1)-\hbar^2m^2\mp\hbar^2m\\&=\hbar^2\{(n\mp m)(n\pm m+1)\},\end{aligned}$$

where we used the fact that $(A_\pm)^*=A_\mp$ and equations (25.14) for $A_\mp A_\pm$. Taking the square root, one finds that

$$\mathcal{N}_\pm(q, n, m)=e^{i\alpha}\hbar\{(n\mp m)(n\pm m+1)\}^{\frac{1}{2}}$$

and in the arbitrary phase factor $e^{i\alpha}$ we have taken $\alpha=0$.

VI.3. We have, from equation (25.31), that

$$\mathbf{n}\boldsymbol{\sigma}=\sum_{r=1}^{3}n_r\sigma_r=\begin{bmatrix}n_3, & n_1-in_2\\ n_1+in_2, & -n_3\end{bmatrix}.$$

VI.5. The general 2×2 Hermitian matrix can be written in terms of four real numbers a, b, c and d as

$$\begin{bmatrix}a & b+ic\\ b-ic & d\end{bmatrix}=\begin{bmatrix}\frac{a+d}{2}+\frac{a-d}{2} & b+ic\\ b-ic & \frac{a+d}{2}-\frac{a-d}{2}\end{bmatrix}$$

$$=\frac{a+d}{2}\mathbf{I}+\frac{a-d}{2}\sigma_3+b\sigma_1-c\sigma_2.$$

VI.6. (*i*) Since the direction of the z-axis is in no way preferred, it follows that the component of $\boldsymbol{\sigma}$ in any direction has the eigenvalues ± 1. We must find two eigenfunctions u_1 and u_2 such that

$$\sigma_n u_1 = u_1, \quad \sigma_n u_2 = -u_2. \tag{H 21}$$

Since α and β form a basis in $\mathfrak{S}$, u_1 and u_2 must be linear combinations of α and β,

$$u_1 = A\alpha + B\beta, \quad u_2 = C\alpha + D\beta \tag{H 22}$$

and

$$|A|^2 + |B|^2 = 1, \quad |C|^2 + |D|^2 = 1 \tag{H 23}$$

so that u_1 and u_2 are normed.

From (25.31) and (26.7) one finds that

$$\begin{array}{lll} \sigma_x\alpha = \beta & \sigma_y\alpha = i\beta & \sigma_z\alpha = \alpha \\ \sigma_x\beta = \alpha & \sigma_y\beta = -i\alpha & \sigma_z\beta = -\beta \end{array}$$

and since

$$\sigma_n = \boldsymbol{\sigma}\mathbf{n} = \sigma_x \sin\theta\cos\varphi + \sigma_y \sin\theta\sin\varphi + \sigma_z\cos\theta$$

the eigenvalue equation for u_1, (H 21), becomes

$$(A\beta + B\alpha)\sin\theta\cos\varphi + i(A\beta - B\alpha)\sin\theta\sin\varphi + (A\alpha - B\beta)\cos\theta = A\alpha + B\beta,$$

and taking the inner products with α and β in turn we get

$$\begin{aligned} B\sin\theta\, e^{-i\varphi} + A\cos\theta &= A \\ A\sin\theta\, e^{i\varphi} - B\cos\theta &= B. \end{aligned}$$

One can easily check that these two equations are consistent, *i.e.* A/B has the same value calculated from either of them. From the first equation, one finds that

$$\begin{aligned} B\sin\theta\, e^{-i\varphi} &= A(1-\cos\theta) \\ B &= A\tan(\theta/2)e^{i\varphi}, \end{aligned}$$

and from (H 23)

$$|A| = \cos\frac{\theta}{2}.$$

We choose the phase factor in A as unity, *i.e.*

$$A = \cos\frac{\theta}{2}$$

so that

$$B = \sin\frac{\theta}{2}e^{i\varphi},$$

and

$$u_1 = \alpha\cos\frac{\theta}{2} + \beta\sin\frac{\theta}{2}e^{i\varphi}. \tag{H 24}$$

We could similarly find u_2, but we determine it more easily from the fact that it must be orthogonal to u_1, *i.e.*

$$(u_1, u_2) = 0,$$

which, with suitable choice of phase factor in u_2, easily gives

$$u_2 = \alpha \sin\frac{\theta}{2} - \beta \cos\frac{\theta}{2} e^{i\varphi}. \tag{H 25}$$

The above calculation admits a simple interpretation. Starting with a coordinate system (x, y, z) we have rotated it about the origin to form a new system (x', y', z') so that the new z-axis, *i.e.* the z'-axis, points in the direction of the unit vector $\mathbf{n}$. The component of the spin vector $\boldsymbol{\sigma}$ in the direction of the z'-axis is $\sigma_{z'} = \sigma_n$ and u_1 and u_2 are its eigenfunctions.

(*ii*) To find the probability of observing $\pm\hbar/2$ when measuring $s_n = \hbar\sigma_n/2$ on a system in a state

$$\Psi = \psi_1(\mathbf{r})\,\alpha + \psi_2(\mathbf{r})\,\beta$$

we must expand this state in eigenfunctions of σ_n, *i.e.* since

$$\alpha = u_1 \cos\frac{\theta}{2} + u_2 \sin\frac{\theta}{2}, \quad \beta = \left(u_1 \sin\frac{\theta}{2} - u_2 \cos\frac{\theta}{2}\right) e^{-i\varphi},$$

$$\Psi = u_1\left\{\psi_1 \cos\frac{\theta}{2} + \psi_2 \sin\frac{\theta}{2} e^{-i\varphi}\right\} + u_2\left\{\psi_1 \sin\frac{\theta}{2} - \psi_2 \cos\frac{\theta}{2} e^{-i\varphi}\right\}.$$

The probability of measuring $+\hbar/2$ for s_n, irrespective of the position of the electron, is then

$$\left(\psi_1 \cos\frac{\theta}{2} + \psi_2 \sin\frac{\theta}{2} e^{-i\varphi}, \quad \psi_1 \cos\frac{\theta}{2} + \psi_2 \sin\frac{\theta}{2} e^{-i\varphi}\right)$$
$$= \cos^2\frac{\theta}{2}(\psi_1, \psi_1) + \sin^2\frac{\theta}{2}(\psi_2, \psi_2) + \sin\theta \operatorname{Re}\left[(\psi_1, \psi_2)e^{-i\varphi}\right],$$

where $\operatorname{Re}(z)$ denotes the real part of the complex number $z = x + iy$, *i.e.* $\operatorname{Re}(z) = x$.

One similarly finds the probability of measuring $-\hbar/2$ to be

$$\sin^2\frac{\theta}{2}(\psi_1, \psi_1) + \cos^2\frac{\theta}{2}(\psi_2, \psi_2) - \sin\theta \operatorname{Re}\left[(\psi_1, \psi_2)e^{-i\varphi}\right].$$

These two probabilities add up to unity as required.

(*iii*) We do not know the wave function of the system completely but only that it is of the form

$$\Phi = c_1\alpha + c_2 e^{i\delta}\beta,$$

where $e^{i\delta}$ represents an undetermined phase factor. Though multiplying the wave function as a whole by a phase factor does not lead to a physically new situation, multiplying part of a wave function by a phase factor does lead to observable differences. Carrying out the same analysis for Φ which we did for Ψ in (*ii*), we find for the probability of s_n having the value $\pm\hbar/2$,

$$|c_1|^2\begin{Bmatrix}\cos^2\\ \sin^2\end{Bmatrix}\frac{\theta}{2} + |c_2|^2\begin{Bmatrix}\sin^2\\ \cos^2\end{Bmatrix}\frac{\theta}{2} \pm \sin\theta\, Re\left[c_1\bar{c}_2 e^{i(\varphi-\delta)}\right].$$

We must now average over the unknown phase δ which can assume all values $0 \leqslant \delta < 2\pi$. Since

$$\frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-i\delta} d\delta = 0,$$

the probabilities, after averaging over δ, become

$$|c_1|^2 \begin{Bmatrix} \cos^2 \\ \sin^2 \end{Bmatrix} \frac{\theta}{2} + |c_2|^2 \begin{Bmatrix} \sin^2 \\ \cos^2 \end{Bmatrix} \frac{\theta}{2}.$$

Cases (*ii*) and (*iii*) show an interesting difference. In case (*iii*) we are dealing with a problem which is axially symmetric about the z-axis. Hence the azimuthal angle φ cannot occur in any physically observable quantities, such as the above probabilities.

The wave-function Ψ in case (*ii*), on the other hand, is not axially symmetric on account of the phase relationship between the two parts $\psi_1 \alpha$ and $\psi_2 \beta$ of the wave function Ψ, which we assumed known. In a symmetric case we would have $\langle s_x \rangle = 0$; for Ψ one easily finds that

$$\begin{aligned} \langle s_x \rangle &= \frac{\hbar}{2}([\psi_1 \alpha + \psi_2 \beta], \sigma_x[\psi_1 \alpha + \psi_2 \beta]) \\ &= \frac{\hbar}{2}(\psi_1 \alpha + \psi_2 \beta, \psi_1 \beta + \psi_2 \alpha) = \hbar Re[(\psi_1, \psi_2)] \neq 0, \end{aligned}$$

showing at once that there is no symmetry about the z-axis. Consequently physically observable results, such as the probabilities we calculated, will in case (*ii*) depend on both θ and φ. Thus, a knowledge of the wave functions ψ_1 and ψ_2 gives one much more information about the system than a knowledge of $|c_1|$ and $|c_2|$ only.

VI.7. From equations (27.3) and (27.4), it follows that

$$\begin{aligned} [M_x, M_y] &= [M_x(1) + M_x(2), M_y(1) + M_y(2)] \\ &= [M_x(1), M_y(1)] + [M_x(2), M_y(2)] \\ &= i\hbar\{M_z(1) + M_z(2)\} \\ &= i\hbar M_z, \ etc. \end{aligned}$$

Since

$$M^2 = M^2(1) + M^2(2) + 2\mathbf{M}(1)\mathbf{M}(2),$$

and from (27.3) and (25.7)

$$[M^2(k), M_x(l)] = 0, \quad (k, l = 1, 2)$$

etc, it follows that

$$[M^2, M_x] = 2[\mathbf{M}(1)\mathbf{M}(2), M_x(1) + M_x(2)];$$

and

$$\begin{aligned} [\mathbf{M}(2)\mathbf{M}(1), M_x(1)] &= M_y(2)[M_y(1), M_x(1)] + M_z(2)[M_z(1), M_x(1)] \\ &= i\hbar\{-M_y(2)M_z(1) + M_z(2)M_y(1)\}, \end{aligned}$$

and similarly

$$[\mathbf{M}(1)\mathbf{M}(2), M_x(2)] = i\hbar\{-M_y(1)M_z(2) + M_z(1)M_y(2)\},$$

so that

$$[M^2, M_x] = 0,$$

etc, proving relations (25.6–7). We need not derive equations (25.12–14) anew, as in obtaining these, equations (25.6–7) only were used.

VI.8. We are combining the three orbital angular momentum eigenfunctions of a p-state ($l = 1$), $\Psi_1(1, m_l)$, ($m_l = -1, 0, 1$) with the spin states $\Psi_2(\frac{1}{2}, -\frac{1}{2}) = \beta$ and $\Psi_2(\frac{1}{2}, \frac{1}{2}) = \alpha$. The resultant 6-dimensional product manifold $\mathfrak{M}_1(1) \times \mathfrak{S}_2$ decomposes into a 4-dimensional manifold $\mathfrak{M}_{\frac{3}{2}}$ corresponding to $j = 3/2$ and a 2-dimensional manifold $\mathfrak{M}_{\frac{1}{2}}$ corresponding to $j = \frac{1}{2}$.

Using the methods of section 27, we obtain for the eigenfunctions $\Psi(1, \frac{1}{2}; j, j_z)$:

(a) $\mathfrak{M}_{\frac{3}{2}}$:

$$\Psi(1, \tfrac{1}{2}; \tfrac{3}{2}, \tfrac{3}{2}) = \Psi_1(1, 1)\Psi_2(\tfrac{1}{2}, \tfrac{1}{2}) = \Psi_1(1, 1)\alpha, \tag{H 26a}$$

$$\begin{aligned}\Psi(1, \tfrac{1}{2}; \tfrac{3}{2}, \tfrac{1}{2}) &= \frac{1}{\sqrt{3}}\{\sqrt{2}\Psi_1(1, 0)\Psi_2(\tfrac{1}{2}, \tfrac{1}{2}) + \Psi_1(1, 1)\Psi_2(\tfrac{1}{2}, -\tfrac{1}{2})\}\\ &= \frac{1}{\sqrt{3}}\{\sqrt{2}\Psi_1(1, 0)\alpha + \Psi_1(1, 1)\beta\}\,;\end{aligned} \tag{H 26b}$$

$$\begin{aligned}\Psi(1, \tfrac{1}{2}; \tfrac{3}{2}, -\tfrac{1}{2}) &= \frac{1}{\sqrt{3}}\{\sqrt{2}\Psi_1(1, 0)\Psi_2(\tfrac{1}{2}, -\tfrac{1}{2}) + \Psi_1(1, -1)\Psi_2(\tfrac{1}{2}, \tfrac{1}{2})\}\\ &= \frac{1}{\sqrt{3}}\{\sqrt{2}\Psi_1(1, 0)\beta + \Psi_1(1, -1)\alpha\}\,;\end{aligned} \tag{H 26c}$$

$$\Psi(1, \tfrac{1}{2}; \tfrac{3}{2}, -\tfrac{3}{2}) = \Psi_1(1, -1)\Psi_2(\tfrac{1}{2}, -\tfrac{1}{2}) = \Psi_1(1, -1)\beta. \tag{H 26d}$$

(b) $\mathfrak{M}_{\frac{1}{2}}$: We obtain $\Psi(1, \frac{1}{2}; \frac{1}{2}, \frac{1}{2})$ by choosing a linear combination of $\Psi_1(1, 0)\alpha$ and $\Psi_1(1, 1)\beta$ orthogonal to $\Psi(1, \frac{1}{2}; \frac{3}{2}, \frac{1}{2})$, equation (H 26b), or so that A_+ operating on it vanishes:

$$A_+\{a\Psi_1(1, 0)\Psi_2(\tfrac{1}{2}, \tfrac{1}{2}) + b\Psi_1(1, 1)\Psi_2(\tfrac{1}{2}, -\tfrac{1}{2})\} = 0$$

i.e.

$$a\sqrt{2} + b = 0;$$

hence

$$\begin{aligned}\Psi(1, \tfrac{1}{2}; \tfrac{1}{2}, \tfrac{1}{2}) &= \frac{1}{\sqrt{3}}\{\Psi_1(1, 0)\Psi_2(\tfrac{1}{2}, \tfrac{1}{2}) - \sqrt{2}\Psi_1(1, 1)\Psi_2(\tfrac{1}{2}, -\tfrac{1}{2})\}\\ &= \frac{1}{\sqrt{3}}\{\Psi_1(1, 0)\alpha - \sqrt{2}\Psi_1(1, 1)\beta\}\end{aligned} \tag{H 27a}$$

$$\begin{aligned}\Psi(1, \tfrac{1}{2}; \tfrac{1}{2}, -\tfrac{1}{2}) &= \frac{1}{\sqrt{3}}\{-\Psi_1(1, 0)\Psi_2(\tfrac{1}{2}, -\tfrac{1}{2}) + \sqrt{2}\Psi_1(1, -1)\Psi_2(\tfrac{1}{2}, \tfrac{1}{2})\}\\ &= \frac{1}{\sqrt{3}}\{-\Psi_1(1, 0)\beta + \sqrt{2}\Psi_1(1, -1)\alpha\}.\end{aligned} \tag{H 27b}$$

VI.9. To see how the 8-dimensional spin space $\mathfrak{S}_1 \times \mathfrak{S}_2 \times \mathfrak{S}_3$ decomposes into manifolds corresponding to given eigenvalues of S^2, we first combine $\mathfrak{S}_1$ and $\mathfrak{S}_2$ according to the scheme of Example 2, section 27:

$$\mathfrak{S}_1 \times \mathfrak{S}_2 = \mathfrak{M}_1(12) + \mathfrak{M}_0(12)$$

(where we have written $\mathfrak{M}_1(12)$ and $\mathfrak{M}_0(12)$ to indicate that these are product spaces of particles 1 and 2), and then we combine each of these with $\mathfrak{S}_3$, according to the schemes

$$\mathfrak{M}_1(12) \times \mathfrak{S}_3 = \mathfrak{M}_{\frac{3}{2}} + \mathfrak{M}_{\frac{1}{2}}' \tag{H 28a}$$

$$\mathfrak{M}_0(12) \times \mathfrak{S}_3 = \mathfrak{M}_{\frac{1}{2}}'', \tag{H 28b}$$

where we have used primes to distinguish the two doublet manifolds.

We span the manifold $\mathfrak{M}_1(12)$ by the eigenfunctions (27.16a–c) and $\mathfrak{M}_0(12)$ by (27.16d).

In the decomposition (H 28a) we then combine an angular momentum 1 with an angular momentum $\frac{1}{2}$. This is exactly the problem we treated in Exercise VI.8, and the results we obtained in that exercise hold here. For in considering the addition of angular momenta in section 27 we were not interested in what the actual wave functions are, but only that they should satisfy certain eigenvalue equations. Hence we obtain the eigenfunctions $\Psi(1, \frac{1}{2}; \frac{3}{2}, m)$ and $\Psi(1, \frac{1}{2}; \frac{1}{2}, m)$ spanning the manifolds $\mathfrak{M}_{\frac{3}{2}}$ and $\mathfrak{M}_{\frac{1}{2}}'$ in (H 28a) by combining equations (H 26a–d), (H 27a–b) and (27.16a–c):

(*a*) $\mathfrak{M}_{\frac{3}{2}}$:

$$\Psi(\tfrac{3}{2}, \tfrac{3}{2}) = \alpha_1\alpha_2\alpha_3 \tag{H 29a}$$

$$\Psi(\tfrac{3}{2}, \tfrac{1}{2}) = \frac{1}{\sqrt{3}}\{\alpha_1\alpha_2\beta_3 + \alpha_2\alpha_3\beta_1 + \alpha_3\alpha_1\beta_2\} \tag{H 29b}$$

$$\Psi(\tfrac{3}{2}, -\tfrac{1}{2}) = \frac{1}{\sqrt{3}}\{\alpha_1\beta_2\beta_3 + \alpha_2\beta_3\beta_1 + \alpha_3\beta_1\beta_2\} \tag{H 29c}$$

$$\Psi(\tfrac{3}{2}, -\tfrac{3}{2}) = \beta_1\beta_2\beta_3; \tag{H 29d}$$

(*b*) $\mathfrak{M}_{\frac{1}{2}}'$:

$$\Psi(\tfrac{1}{2}, \tfrac{1}{2}) = \frac{1}{\sqrt{3}}\left\{\frac{1}{\sqrt{2}}(\alpha_1\beta_2 + \alpha_2\beta_1)\alpha_3 - \sqrt{2}\alpha_1\alpha_2\beta_3\right\} \tag{H 30a}$$

$$\Psi(\tfrac{1}{2}, -\tfrac{1}{2}) = \frac{1}{\sqrt{3}}\left\{-\frac{1}{\sqrt{2}}(\alpha_1\beta_2 + \alpha_2\beta_1)\beta_3 + \sqrt{2}\beta_1\beta_2\alpha_3\right\}. \tag{H 30b}$$

We easily obtain the other 2-dimensional manifold $\mathfrak{M}_{\frac{1}{2}}''$ arising from combining the singlet state (27.16d)

$$\frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \alpha_2\beta_1)$$

with α_3 and β_3:

(*c*) $\mathfrak{M}_{\frac{1}{2}}''$:

$$\Psi(\tfrac{1}{2}, \tfrac{1}{2}) = \frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \alpha_2\beta_1)\alpha_3 \qquad \text{(H 31a)}$$

$$\Psi(\tfrac{1}{2}, -\tfrac{1}{2}) = \frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \alpha_2\beta_1)\beta_3. \qquad \text{(H 31b)}$$

If we again consider the symmetry under the permutation of particles, it follows from the fact that the triplet states of particles 1 and 2 are symmetric and that $\mathfrak{M}_{\frac{3}{2}}$ and $\mathfrak{M}_{\frac{1}{2}}'$ are derived from these, that $\mathfrak{M}_{\frac{3}{2}}$ and $\mathfrak{M}_{\frac{1}{2}}'$ contain only wave functions symmetric in 1 and 2. Furthermore, $\mathfrak{M}_{\frac{3}{2}}$ contains only wave functions symmetric in all three particles, as one sees from equations (H 29a–d). Similarly $\mathfrak{M}_{\frac{1}{2}}''$ contains only wave functions antisymmetric in particles 1 and 2 as it is derived from the singlet state $\frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \alpha_2\beta_1)$.

We could of course equally well have combined the spin states of any other pair of particles to begin with. This corresponds to permuting the suffixes 1, 2 and 3, labelling the particles, in the above equations. This does not affect $\mathfrak{M}_{\frac{3}{2}}$ which is completely symmetric, but the resultant 2-dimensional manifolds $\mathfrak{M}_{\frac{1}{2}}'$ and $\mathfrak{M}_{\frac{1}{2}}''$ are different. Together they, of course, span the same 4-dimensional manifold

$$\mathfrak{S}_1 \times \mathfrak{S}_2 \times \mathfrak{S}_3 - \mathfrak{M}_{\frac{3}{2}}.$$

In practice, this ambiguity of states does not usually arise. On account of the Pauli principle (section 28), the symmetry properties of the wave functions are usually prescribed.

VI.10. By simply interchanging the labelling 1 and 2 of the particles we have $H(2, 1)\psi(2, 1) = E\psi(2, 1)$, and since $H(1, 2) = H(2, 1)$ the result follows.

VI.11. Substituting (28.3) into (28.4) gives

$$(|a|^2 - |b|^2)|\psi(1, 2)|^2 + (|b|^2 - |a|^2)|\psi(2, 1)|^2 \\ + (a\bar{b} - b\bar{a})\psi(1, 2)\overline{\psi(2, 1)} + (b\bar{a} - a\bar{b})\psi(2, 1)\overline{\psi(1, 2)} = 0.$$

Since $\psi(1, 2)$ is an arbitrary wave function, this equation is only satisfied if

$$|a|^2 = |b|^2 \quad \text{and} \quad a\bar{b} = b\bar{a}.$$

Hence, if $a = A\mathrm{e}^{i\alpha}$ and $b = B\mathrm{e}^{i\beta}$ (A, B, α, β real and $A > 0$, $B > 0$), we require

$$A = B \qquad \mathrm{e}^{i(\alpha-\beta)} = \mathrm{e}^{i(\beta-\alpha)},$$

i.e. $(\alpha - \beta) = n\pi$, $(n = 0, 1, 2, \ldots)$, so that $a = \pm b$, and (28.5) follows.

VI.12. Let 1 and 2 denote the protons, and 3 the neutron. There are two doublet spin states which we obtained in Exercise VI.9, equations (H 30) and (H 31); we write them

$$\xi_a^+(\widetilde{12};3)=\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1)\alpha_3$$

$$\xi_a^-(\widetilde{12};3)=\frac{1}{\sqrt{2}}(\alpha_1\beta_2-\alpha_2\beta_1)\beta_3$$

and

$$\xi_s^+(\overline{12};3)=\frac{1}{\sqrt{3}}\left\{\frac{1}{\sqrt{2}}(\alpha_1\beta_2+\alpha_2\beta_1)\alpha_3-\sqrt{2}\alpha_1\alpha_2\beta_3\right\}$$

$$\xi_s^-(\overline{12};3)=\frac{1}{\sqrt{3}}\left\{-\frac{1}{\sqrt{2}}(\alpha_1\beta_2+\alpha_2\beta_1)\beta_3+\sqrt{2}\beta_1\beta_2\alpha_3\right\}.$$

The total wave function must be antisymmetric in the proton coordinates, but there is no restriction as far as the single neutron is concerned. If $\varphi_s(\overline{12};3)$ and $\varphi_a(\widetilde{12};3)$ denote spatial wave functions of the indicated symmetry properties then the total wave function is of the form

$$\Phi^\pm(\widetilde{12};3)=a\varphi_s(\overline{12};3)\xi_a^\pm(\widetilde{12};3)+b\varphi_a(\widetilde{12};3)\xi_s^\pm(\overline{12};3),$$

where the ± signs indicate the spin states with $S_z=\pm\hbar/2$. As for the α-particle, one finds that $|b|\ll|a|$; *i.e.* the two protons have a great tendency to be in a singlet state. (*Cf* end of section 28.)

CHAPTER VII

VII.2. The perturbation of any energy level E_n, equation (11.15), is given by equation (29.13).

$$V_{nn}=a\int_{-\infty}^{\infty}x^3|\psi_n(x)|^2dx.$$

Since the perturbation $V(x)=ax^3$ is an odd function, $V(-x)=-V(x)$, and the oscillator wave function $\psi_n(x)$, equation (11.23), has always a definite parity, so that $|\psi_n(-x)|^2=|\psi_n(x)|^2$, it follows that

$$V_{nn}=0.$$

This result is generally true for states of definite parity and perturbations of odd parity. Second-order perturbation theory will, in general, give a displacement of levels.

VII.3. Using the notation of section 11, the perturbed energy levels are, according to (11.15) and (29.13)

$$E_n=\left(n+\frac{1}{2}\right)\hbar\omega_c+V_{nn},$$

where

$$V_{nn}=a\int_{-\infty}^{\infty}\{x^2\psi_n(x)\}^2dx=\frac{a}{a^5}\int_{-\infty}^{\infty}\{\xi^2\varphi_n(\xi)\}^2d\xi,$$

where we have used the fact that the oscillator eigenfunctions, equation (11.23), are real functions. Applying the recurrence relation (11.19) twice, one obtains

$$\xi^2H_n(\xi)=\tfrac{1}{4}\,H_{n+2}(\xi)+\tfrac{1}{2}(2n+1)H_n(\xi)+n(n-1)H_{n-2}(\xi),$$

and using the orthogonality relation (11.22), one obtains

$$V_{nn}=\frac{3}{4}\left(\frac{a\hbar^2}{m^2\omega_0^2}\right)(2n^2+2n+1).$$

VII.4. To carry out the second-order perturbation calculation, we must solve equation (29.8c), as well as (29.8a–b). Taking equations (29.5) to (29.14) over unchanged and putting

$$\varphi_{\mathcal{N}}^{(2)}=\sum_{r=1}^{\infty}b_r\varphi_r^{(0)},$$

equation (29.8c) gives on substitution

$$\sum_{r=1}^{\infty}b_rE_r^{(0)}\varphi_r^{(0)}+V\sum_{r=1}^{\infty}a_r\varphi_r^{(0)}=E_{\mathcal{N}}^{(0)}\sum_{r=1}^{\infty}b_r\varphi_r^{(0)}+E_{\mathcal{N}}^{(1)}\sum_{r=1}^{\infty}a_r\varphi_r^{(0)}+E_{\mathcal{N}}^{(2)}\varphi_{\mathcal{N}}^{(0)};$$

and taking inner products with $\varphi_s^{(0)}$:

$$b_sE_s^{(0)}+\sum_{r=1}^{\infty}a_rV_{sr}=b_sE_{\mathcal{N}}^{(0)}+a_sE_{\mathcal{N}}^{(1)}+E_{\mathcal{N}}^{(2)}\delta_{\mathcal{N}s}.$$

Separating the cases $s=\mathcal{N}$ and $s\neq\mathcal{N}$, the last equation gives, using equations (29.13–14),

$$E_{\mathcal{N}}^{(2)}=\sum_r{}'\,\frac{V_{\mathcal{N}r}V_{r\mathcal{N}}}{E_{\mathcal{N}}^{(0)}-E_r^{(0)}},\quad(s=\mathcal{N}),$$

and

$$b_s=\sum_r{}'\,\frac{V_{sr}V_{r\mathcal{N}}}{(E_{\mathcal{N}}^{(0)}-E_s^{(0)})\,(E_{\mathcal{N}}^{(0)}-E_r^{(0)})}-\frac{V_{\mathcal{N}\mathcal{N}}V_{s\mathcal{N}}}{(E_{\mathcal{N}}^{(0)}-E_s^{(0)})^2}+\frac{a_{\mathcal{N}}V_{s\mathcal{N}}}{E_{\mathcal{N}}^{(0)}-E_s^{(0)}},\ (s\neq\mathcal{N})$$

where, as before, the term $r=\mathcal{N}$ is excluded from the summation $\sum'$.

We again determine $a_{\mathcal{N}}$ and $b_{\mathcal{N}}$ from the normalization condition

$$\|\varphi_{\mathcal{N}}^{(0)}+\varepsilon\varphi_{\mathcal{N}}^{(1)}+\varepsilon^2\varphi_{\mathcal{N}}^{(2)}\|=1$$

which must now be satisfied up to terms in ε^2. In this way, we obtain the two conditions

$$a_{\mathcal{N}}+\bar{a}_{\mathcal{N}}=0$$

$$b_{\mathcal{N}}+b_{\mathcal{N}}=-\sum_r|a_r|^2.$$

These two equations again fix the real parts of $a_{\mathcal{N}}$ and $b_{\mathcal{N}}$. As in section 29, we can again put the imaginary parts of $a_{\mathcal{N}}$ and $b_{\mathcal{N}}$ equal to zero, corresponding to a particular choice of phase factor for φ_N, so that

$$a_N = 0$$
$$b_N = -\tfrac{1}{2}\sum_r |a_r|^2.$$

Using these results, we finally obtain the following expression for the energy levels and the energy eigenfunctions according to second-order perturbation theory:

$$E_N = E_N^{(0)} + \varepsilon V_{NN} + \varepsilon^2 \sum{}' \frac{|V_{Nr}|^2}{E_N^{(0)} - E_r^{(0)}}, \tag{H 32}$$

$$\begin{aligned}\varphi_N = \varphi_N^{(0)} &+ \varepsilon \sum_r{}' \varphi_r^{(0)} \frac{V_{rN}}{E_N^{(0)} - E_r^{(0)}} \\ &+ \varepsilon^2 \sum_r{}' \left\{ \varphi_r^{(0)} \left[\sum_s{}' \frac{V_{rs} V_{sN}}{(E_N^{(0)} - E_r^{(0)})(E_N^{(0)} - E_s^{(0)})} - \frac{V_{rN} V_{NN}}{(E_N^{(0)} - E_r^{(0)})^2} \right] \right. \\ &\qquad \left. - \tfrac{1}{2} \varphi_N^{(0)} \frac{|V_{rN}|^2}{(E_N^{(0)} - E_r^{(0)})^2} \right\}. \end{aligned} \tag{H 33}$$

VII.5. The matrix $H = \begin{bmatrix} E_1 & & \\ & E_2 & \\ & & E_3 \end{bmatrix}$ has the eigenvectors $\psi_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, *etc.* Under the perturbation, $H \rightarrow H + \varepsilon V$, $\psi_1 \rightarrow \begin{bmatrix} 1 \\ \lambda \\ \mu \end{bmatrix}$, where λ and μ are of the first order in ε, and $E_1 \rightarrow E_1 + \varepsilon E_1'$. The eigenvalue problem becomes

$$(H + \varepsilon V) \begin{bmatrix} 1 \\ \lambda \\ \mu \end{bmatrix} = (E_1 + \varepsilon E_1') \begin{bmatrix} 1 \\ \lambda \\ \mu \end{bmatrix},$$

i.e. to first order in ε:

$$\begin{bmatrix} 0 \\ E_2\lambda \\ E_3\mu \end{bmatrix} + \varepsilon \begin{bmatrix} V_{11} \\ V_{21} \\ V_{31} \end{bmatrix} = E_1 \begin{bmatrix} 0 \\ \lambda \\ \mu \end{bmatrix} + \varepsilon E_1' \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Hence

$$\left.\begin{aligned} E_1' &= V_{11} \\ \lambda(E_1 - E_2) &= \varepsilon V_{21} \\ \mu(E_1 - E_3) &= \varepsilon V_{31} \end{aligned}\right\},$$

which is just what we should expect from equations (29.13–14).

If $E_1 = E_2$, the above method breaks down: λ is no longer small. This is because any vector $\begin{bmatrix} \lambda \\ \mu \\ 0 \end{bmatrix}$ is an eigenvector of H belonging to the eigenvalue E_1. Hence we now write the corresponding perturbed eigenvector $\begin{bmatrix}\lambda\\ \mu\\ \nu\end{bmatrix}$, where only ν is small, of order ε. To first order in ε, the eigenvalue problem now gives (denoting the perturbed eigenvalue by $E_1+\varepsilon E'$)

$$\left.\begin{aligned}\varepsilon(V_{11}\lambda+V_{12}\mu)&=\varepsilon E'\lambda\\ \varepsilon(V_{21}\lambda+V_{22}\mu)&=\varepsilon E'\mu\\ E_3\nu+\varepsilon(V_{31}\lambda+V_{32}\mu)&=E_1\nu\end{aligned}\right\}.$$

The last equation gives

$$\nu=\frac{\varepsilon(V_{31}\lambda+V_{32}\mu)}{E_1-E_3}.$$

The first two equations are homogenous in λ and μ and only possess a nontrivial solution if

$$\begin{vmatrix}V_{11}-E' & V_{12}\\ V_{21} & V_{22}-E'\end{vmatrix}=0.$$

This is the secular equation which is quadratic in E' and thus gives two solutions E_1' and E_2'. Corresponding to each of these, we obtain a value of λ/μ and hence, apart from normalization, an eigenvector (*cf* sections 22 and 30).

CHAPTER VIII

VIII.1. For a solution depending on the radial coordinate r only, ∇^2 reduces to

$$\nabla^2=\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right).$$

Hence

$$\begin{aligned}(\nabla^2+k^2)\frac{e^{ikr}}{r}&=\frac{1}{r^2}\frac{d}{dr}\left[e^{ikr}\left\{rik+r^2\frac{d}{dr}\left(\frac{1}{r}\right)\right\}\right]+k^2\frac{e^{ikr}}{r}\\ &=\frac{ik}{r^2}e^{ikr}-\frac{k^2}{r}e^{ikr}+ike^{ikr}\frac{d}{dr}\left(\frac{1}{r}\right)\\ &\quad+e^{ikr}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)\frac{1}{r}+\frac{k^2}{r}e^{ikr}\\ &=e^{ikr}\nabla^2\left(\frac{1}{r}\right)\end{aligned}$$

so that

$$(\nabla^2+k^2)\left(-\frac{e^{ikr}}{4\pi r}\right)=e^{ikr}\left(-\frac{1}{4\pi}\nabla^2\frac{1}{r}\right)=e^{ikr}\delta(r), \qquad \text{(H 34)}$$

from equation (33.2). We can omit the exponential factor e^{ikr} from the right-hand term of equation (H 34), since $\delta(r)$ vanishes, except at $r=0$, and in that case the exponential becomes unity, proving the required result.

The result for the second solution $(-e^{-ikr}/4\pi r)$ follows at once by taking complex conjugates throughout the above proof.

VIII.2. From the discussion leading to equation (33.16), we expect the required transition probability per unit time to be

$$\frac{2\pi}{\hbar}\left|\left(e^{i\mathbf{k}_s\mathbf{r}}\varphi_s(q),\ V(\mathbf{r},\ q)\psi(\mathbf{r},\ q)\right)\right|^2\frac{d\mathbf{k}_s}{(2\pi)^3}\delta(E'-E), \qquad \text{(H 35)}$$

where E is the initial energy of the system, equation (36.3), and E' the final energy:

$$E'=E_s+\frac{\hbar^2}{2\mu}k_s^2.$$

One easily verifies this by integrating (H 35) over dk_s to eliminate the δ-function, and dividing by the flux of incident particles, $\hbar k_0/\mu$, to give the same expression for the cross-section as before, equation (36.14).

VIII.3. To obtain the first Born approximation, we replace the true wave function $\psi(\mathbf{r}, q)$ by the zero-order approximation

$$e^{i\mathbf{k}_0\mathbf{r}}\varphi_0(q)$$

[*cf* equation (36.4)] on the right-hand side of equation (36.6). This gives for the function $F_s(\mathbf{r})$, defined in equation (36.8),

$$F_s(\mathbf{r})=e^{i\mathbf{k}_0\mathbf{r}}\int dq\overline{\varphi_s(q)}\,U(\mathbf{r},\ q)\varphi_0(q)\equiv e^{i\mathbf{k}_0\mathbf{r}}\frac{2\mu}{\hbar^2}W(\mathbf{r}).$$

$W(\mathbf{r})$ represents a mean potential, averaged over the internal coordinates of the scatterer, which the incident particle experiences.

The cross-section then becomes, from equations (36.12) and (36.14),

$$\sigma(\theta,\ \varphi)=\left(\frac{k_s}{k_0}\right)\left(\frac{\mu}{2\pi\hbar^2}\right)^2\left|\int W(\mathbf{r})e^{i(\mathbf{k}_0-\mathbf{k}_s)\mathbf{r}}d\mathbf{r}\right|^2,$$

[*cf* equation (34.1)]. $\hbar(\mathbf{k}_0-\mathbf{k}_s)$ again represents the momentum which the incident particle has transferred to the target during the collision.

VIII.4. If $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of the two particles, then the wave function of the two-particle system can be written (*cf* Exercise III.3)

$$\psi(\mathbf{r}_1,\ \mathbf{r}_2)=\Phi\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}\right)\chi(\mathbf{r}) \quad (\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2),$$

the wave function Φ which describes the motion of the centre of mass of the system being symmetric in the coordinates of the two particles (since they have the same mass), and $\chi(\mathbf{r})$ describing the relative motion.

If we treat the two particles 1 and 2 as different, and hence distinguishable, to begin with, then the asymptotic form of the wave function in the C.o.M. system is

$$e^{ikz}+\frac{1}{r}e^{ikr}f(\theta,\ \varphi),$$

(r, θ, φ) being the polar coordinates of the vector $\mathbf{r}$ giving the position of particle 1 relative to particle 2. Since the polar coordinates of the vector $-\mathbf{r}$ (corresponding to interchanging the two particles) are $(r, \pi-\theta, \pi+\varphi)$, [*cf* equation (13.24)] the asymptotic forms of the wave functions, symmetric and antisymmetric in the spatial coordinates respectively, are

$$(\mathrm{e}^{ikz} \pm \mathrm{e}^{-ikz}) + \frac{\mathrm{e}^{ikr}}{r}[f(\theta, \varphi) \pm f(\pi-\theta, \pi+\varphi)].$$

Hence we obtain for the cross-section, by arguments analogous to those in section 32,

$$\sigma(\theta, \varphi) = |f(\theta, \varphi) \pm f(\pi-\theta, \pi+\varphi)|^2.$$

We can also see this result in another way. If particle 1 is scattered through angles $\pi-\theta, \pi+\varphi$, particle 2 flies off in the opposite direction, θ, φ. Hence the scattering amplitude corresponding to this process must also be included in the cross-section $\sigma(\theta, \varphi)$ if the two particles are identical.

VIII.5. A qualitative justification for neglecting exchange effects in the case of fast electrons was given at the end of section 36. Since we are dealing with fast electrons, the validity criterion for the Born approximation is satisfied, so that we shall describe incident and scattered electrons by plane waves.

To calculate the cross-section, we use the method of section 36. Equations (36.12) and (36.8) then give for the case of elastic scattering in the Born approximation

$$f(\theta) = -\frac{1}{4\pi}\int \mathrm{e}^{-i\mathbf{k}\mathbf{r}} d\mathbf{r} \int d\mathbf{s}\overline{\varphi(s)}\frac{2m}{\hbar^2}\left(-\frac{e^2}{r} + \frac{e^2}{|\mathbf{r}-\mathbf{s}|}\right)\varphi(s)\mathrm{e}^{i\mathbf{k}_0\mathbf{r}},$$

for an electron to be scattered from a state with momentum $\hbar\mathbf{k}_0$ to one with momentum $\hbar\mathbf{k}$ (both states having the same kinetic energy $E = \hbar^2 k_0^2/2m = \hbar^2 k^2/2m$), the wave function of the atomic electron in the ground state of the hydrogen atom being

$$\varphi(s) = \frac{1}{\sqrt{\pi}}\alpha^{\frac{3}{2}}\mathrm{e}^{-\alpha s}.$$

($\mathbf{r}$, $\mathbf{s}$ are position vectors of incident and atomic electrons relative to the nucleus.) We can write this expression for the scattering amplitude

$$f(\theta) = -\frac{m}{2\pi\hbar^2}\int V(r)\mathrm{e}^{i(\mathbf{k}_0-\mathbf{k})\mathbf{r}} d\mathbf{r}$$

where

$$V(r) = -\frac{e^2}{r} + \int e^2|\varphi(s)|^2\frac{1}{|\mathbf{r}-\mathbf{s}|}d\mathbf{s}.$$

Comparison with equation (34.1) shows that $V(r)$ represents the mean potential which the hydrogen atom presents to the incident electron. This is physically reasonable. The first term $(-e^2/r)$ represents the

Coulomb field of the nucleus, the second term (the integral) represents the Coulomb field of the atomic electron whose charge density distribution is $-e|\varphi(s)|^2$. Evaluating $V(r)$, we obtain

$$V(r) = -e^2\left(\frac{1}{r}+a\right)e^{-2ar},$$

showing that we are dealing with a spherically symmetric potential as we indicated by writing $V(r)$ rather than $V(\mathbf{r})$.

We can now use equations (34.3) to (34.5) directly, rather than repeat the angular integration, and obtain

$$f(\theta) = \frac{2me^2}{\hbar^2 K}\int_0^\infty \left(\frac{1}{r}+a\right)e^{-2ar}\sin(Kr)r\,dr$$

$$= \frac{2me^2(8a^2+K^2)}{\hbar^2(4a^2+K^2)^2}.$$

Hence we obtain for the differential cross-section

$$\sigma(\theta) = |f(\theta)|^2 = \frac{4m^2e^4(8a^2+K^2)^2}{\hbar^4(4a^2+K^2)^4}, \qquad \text{(H 36)}$$

where $K = 2k\sin(\theta/2)$, equation (34.5). Integrating equation (H 36) over all scattering angles we obtain the total cross-section,

$$\sigma_{\text{tot}} = 2\pi\int_0^\pi \sigma(\theta)\sin\theta\,d\theta$$

$$= 2\pi\int_{K=0}^{2k} \sigma(\theta)\frac{K\,dK}{k^2},$$

so that

$$\sigma_{\text{tot}} = \frac{8\pi m^2e^4}{\hbar^4k^2}\int_0^{2k}\frac{(8a^2+K^2)^2}{(4a^2+K^2)^4}K\,dK$$

$$= \frac{\pi}{3}\left(\frac{me^2}{\hbar^2}\right)^2\frac{(12a^4+18a^2k^2+7k^4)}{a^2(a^2+k^2)^3}.$$

VIII.6. From equations (38.29) and (32.12)

$$\sigma_{\text{tot}} = \frac{1}{k^2}\sum_l\sum_{l'}(2l+1)(2l'+1)e^{i(\delta_l-\delta_{l'})}\sin\delta_l\sin\delta_{l'}\times 2\pi\int_{-1}^{+1}P_l(x)P_{l'}(x)\,dx.$$

From (13.16–17) this integral equals $2\delta_{ll'}/(2l+1)$, giving at once (38.30).

VIII.8. Consider the potential $\lambda V(r)$, with λ a continuously varying parameter: $0<\lambda<1$. The phase shifts δ_l will then be functions of λ.

Correspondingly, we write instead of (38.38):

$$\frac{d^2}{dr^2}v_l(r, \lambda) + \left[k^2 - \lambda U(r) - \frac{l(l+1)}{r^2}\right]v_l(r, \lambda) = 0$$

and a similar equation with $(\lambda + \Delta\lambda)$ instead of λ. Proceeding as in section 38, one finds

$$\frac{d}{d\lambda}\sin \delta_l(\lambda) = -\frac{1}{k}\int_0^\infty U(r)[v_l(r, \lambda)]^2 dr$$

which is always $\leqslant 0$ or $\geqslant 0$, according as $U \geqslant 0$ or $U \leqslant 0$ respectively.

VIII.9. We can solve the wave equation (38.38) explicitly for the case under consideration (*cf.* section 8 and Exercises III.4. and III.5 where similar problems are treated):

$$v_0(r) = A \sin k_i r, \qquad \text{if } r < a,$$

$$v_0(r) = \sin (kr + \delta_0), \quad \text{if } r > a,$$

where k and k_i are related to the bombarding energy E and the potential depth V_0 by

$$E = \frac{\hbar^2 k^2}{2\mu}, \quad (E + V_0) = \frac{\hbar^2 k_i^2}{2\mu}.$$

The two constants A and δ_0 are to be chosen so that v_0 and dv_0/dr are continuous at $r = a$. This gives

$$\tan (ka + \delta_0) = \frac{k}{k_i}\tan (k_i a),$$

yielding δ_0. As E, *i.e.* k tends to zero, $\delta_0 \to 0$. The corresponding cross-section is given by $\sigma_0 = (4\pi \sin^2 \delta_0)/k^2$. The last equation gives

$$\sigma_0 = 4\pi a^2\left[\frac{\tan k_0 a}{k_0 a} - 1\right]^2,$$

since

$$\operatorname*{Lim}_{k\to 0} k_i = (2\mu V_0/\hbar^2)^{\frac{1}{2}} \equiv k_0.$$

CHAPTER IX

IX.1. For $n = 2$, $\mathfrak{S}_2$ consists of only two elements, the identity E and the permutation P_{12}, and hence is Abelian. (The same argument proves that the inversion group is Abelian.)

For $n > 2$, we need only consider $\mathfrak{S}_3$. For the permutations of $\mathfrak{S}_3$ all occur in $\mathfrak{S}_n$ and hence if these are not commutative, $\mathfrak{S}_n$ cannot be Abelian.

Given three objects A, B, C arranged in definite order, let P_{ij} be the operator which exchanges the objects in the ith and jth places. Then

$$P_{12}P_{13}(ABC)=P_{12}(CBA)=(BCA)$$
$$P_{13}P_{12}(ABC)=P_{13}(BAC)=(CAB),$$

so that $\mathfrak{S}_3$, and hence $\mathfrak{S}_n$ with $n>3$ as well, are not Abelian.

IX.2. Carrying out the permutation is simply a relabelling of the variables of integration in the integral defining the scalar product, and this is hence invariant. We illustrate it for $n=2$:

$$(f(\mathbf{r}_1, \mathbf{r}_2), g(\mathbf{r}_1, \mathbf{r}_2))=(f(\mathbf{r}_2, \mathbf{r}_1), g(\mathbf{r}_2, \mathbf{r}_1))=(P_{12}f(\mathbf{r}_1, \mathbf{r}_2), P_{12}g(\mathbf{r}_1, \mathbf{r}_2)).$$

IX.3. The proofs are closely analogous to that of section 41 for spinless particles. We shall indicate the proof of one of the three results only, namely for the operators $\{S, T, \dots\}$ of the total function space.

From the first of equations (42.6) and (42.8), it follows that for any wave function $\{\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta\}$

$$\begin{aligned}
T\Gamma(\mathbf{r}, \boldsymbol{\sigma})\,&\{\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta\}\\
&=\Gamma(t^{-1}\mathbf{r}, t^{-1}\boldsymbol{\sigma})\,\{\varphi_1(t^{-1}\mathbf{r})[c_1\alpha+c_2\beta]+\varphi_2(t^{-1}\mathbf{r})[c_3\alpha+c_4\beta]\}\\
&=\Gamma(\mathbf{r}, \boldsymbol{\sigma})\,\{\varphi_1(t^{-1}\mathbf{r})[c_1\alpha+c_2\beta]+\varphi_2(t^{-1}\mathbf{r})[c_3\alpha+c_4\beta]\}\\
&=\Gamma(\mathbf{r}, \boldsymbol{\sigma})\,T\{\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta\}.
\end{aligned}$$

The converse follows exactly as equation (41.4) if we define the inverse operator T^{-1} appropriately. From equation (42.4a) this definition is

$$T^{-1}[\varphi_1(\mathbf{r})\alpha+\varphi_2(\mathbf{r})\beta]=\varphi_1(t\mathbf{r})[c_1'\alpha+c_2'\beta]+\varphi_2(t\mathbf{r})[c_3'\alpha+c_4'\beta]$$

where the c_j' are the coefficients of the solution of (42.2) for α' and β':

$$\alpha'=c_1'\alpha+c_2'\beta,\quad \beta'=c_3'\alpha+c_4'\beta.$$

IX.4. For two particles S is a good quantum number. So we still have a pure triplet state. L is no longer a good quantum number, though $J=1$ is. The L values which can combine with $S=1$ to give $J=1$ are $L=0, 1, 2$. However, parity is a good quantum number, and for two particles with relative orbital angular momentum L the parity is $(-1)^L$; *i.e.* the S and D states have opposite parity to the P state. Hence the ground state becomes a mixture of ${}^3S_1$ and ${}^3D_1$ states when tensor forces are present.

IX.5. If ψ_1 is *any* vector in $\mathfrak{M}_1$, then $T^{-1}\psi_1$ is in $\mathfrak{M}_1$ (since $\mathfrak{M}_1$ is invariant under $\mathfrak{G}$). Hence, if ψ_2 is any vector in $\mathfrak{M}_2$, then $(T^{-1}\psi_1, \psi_2)=0$, *i.e.* $(\psi_1, T\psi_2)=0$. Hence $T\psi_2$ is orthogonal to ψ_1, and since ψ_1 is *any* vector of $\mathfrak{M}_1$, $T\psi_2$ is orthogonal to the *whole* space $\mathfrak{M}_1$. But $\mathfrak{M}$ is invariant and ψ_2 a vector of $\mathfrak{M}$, hence $T\psi_2$ lies in $\mathfrak{M}$ and therefore in $\mathfrak{M}_2$.

IX.6. The theorem is proved, if we can prove that any invariant manifold $\mathfrak{M}$ of the group $\mathfrak{G}=\{S, T, \dots\}$ is completely reducible into one-dimensional invariant manifolds.

Let $\mathfrak{M}$ be such an n-dimensional manifold. Since the operators of $\mathfrak{G}$ all commute with each other, they possess a set of orthonormal mutual eigenvectors $\varphi_1, \varphi_2, \dots$ spanning the whole function space $\mathfrak{H}$:

$$S\varphi_i=s_i\varphi_i,\ T\varphi_i=t_i\varphi_i, \dots, \quad (i=1, 2, \dots).$$

We can decompose each vector φ_i uniquely into two parts

$$\varphi_i=\xi_i+\eta_i$$

such that ξ_i and η_i lie in the invariant manifolds $\mathfrak{M}$ and $\mathfrak{H}-\mathfrak{M}$ (*cf* Exercise IX.5) respectively. Hence the equation

$$T\xi_i=t_i\xi_i+(t_i-T)\eta_i$$

implies that

$$T\xi_i=t_i\xi_i, \quad T\eta_i=t_i\eta_i$$

(since $T\xi_i$ must lie in $\mathfrak{M}$, and $(t_i-T)\eta_i$ is orthogonal to $\mathfrak{M}$), *i.e.* the manifolds $\mathfrak{M}$ and $\mathfrak{H}-\mathfrak{M}$ are separately spanned by orthonormal sets of mutual eigenvectors of the group $\mathfrak{G}$, and each such vector defines a one-dimensional invariant subspace.

IX.7. $\mathbf{l}$ and $\boldsymbol{\sigma}$ are pseudo-vectors under total rotations and inversion. Hence we at once find the selection rules:

(*i*) no change of parity;

(*ii*) $\Delta J\equiv J'-J=0, \pm 1$ (not $0\rightarrow 0$).

Let us next again assume that total spin and orbital angular momenta by themselves are good quantum numbers. The orbital and spin parts lead to different selection rules for L and S. $\boldsymbol{\sigma}$ and $\mathbf{l}$ are of course invariant as far as rotations in orbital and spin space respectively are concerned. Hence we obtain these selection rules:

(*a*) orbital part of interaction:

$$\Delta S=0$$
$$\Delta L=0, \pm 1, \quad (\text{not } 0\rightarrow 0);$$

(*b*) spin part of interaction:

$$\Delta S=0, \pm 1, \quad (\text{not } 0\rightarrow 0).$$
$$\Delta L=0.$$

IX.8. We must show that

$$-i\hbar e\int\overline{\varphi'}\,\frac{\partial}{\partial x}\,\varphi d\mathbf{r} \tag{H 37}$$

is proportional to (45.19). If E and E' are the energies of the states φ and φ' then

$$\nabla^2\varphi = -\frac{2m}{\hbar^2}(E-V)\varphi$$

$$\nabla^2\overline{\varphi'} = -\frac{2m}{\hbar^2}(E'-V)\overline{\varphi'},$$

and hence

$$\int d\mathbf{r}(\overline{\varphi'}x\nabla^2\varphi - \varphi x\nabla^2\overline{\varphi'}) = \frac{2m}{\hbar^2}(E'-E)\int\overline{\varphi'}x\varphi d\mathbf{r}. \qquad \text{(H 38)}$$

The right-hand term is proportional to the electric dipole term (45.19). We need only show that the left-hand term is proportional to equation (H 37). We can replace the left-hand term in (H 38) by

$$\int d\mathbf{r}\,\{\overline{\varphi'}\nabla^2(x\varphi) - (x\varphi)\nabla^2\overline{\varphi'}\} - 2\int\overline{\varphi'}\frac{\partial\varphi}{\partial x}d\mathbf{r}. \qquad \text{(H 39)}$$

(This follows, since one easily proves from equations (A4), (A5) and (A6) in Appendix A, that

$$\nabla^2(x\varphi) = x\nabla^2\varphi + 2\frac{\partial\varphi}{\partial x}.\Big)$$

The first term in (H 39) can be transformed into a surface integral, by Green's theorem, equation (A11), over a surface at infinity, which vanishes. The second term is proportional to (H 37), proving our result.

IX.9. The operator

$$\sum_{k=1}^{Z}(z_k p_x(k) + x_k p_z(k))$$

occurring in the matrix element of the quadrupole transitions has parity $(+1)$, and it was so chosen as to transform according to $\mathfrak{D}_2$ under rotations. Since

$$\mathfrak{D}_2 \times \mathfrak{D}_J = \begin{cases} \mathfrak{D}_2, & \text{if } J=0 \\ \mathfrak{D}_{2+J} + \ldots + \mathfrak{D}_{|2-J|} & \text{if } J \geq \frac{1}{2}, \end{cases}$$

we at once have these selection rules (with the same notation as in section 45):

(*i*) $\Delta J = 0, \pm 1, \pm 2;$ $\begin{pmatrix}\text{not } 0 \to 0 \\ \text{not } 0 \to 1 \\ \text{not } 1 \to 0\end{pmatrix}$

(*ii*) no change in parity.

If L and S are also, separately, good quantum numbers, one obtains the additional selection rules:

(*iii*) $\Delta S = 0;$

(*iv*) $\Delta L = \Delta J$ [as under (*i*)]

CHAPTER X

X.2.

$$a_k = \begin{bmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{bmatrix}$$

X.3. $\hat{\sigma}_k = -i\gamma_1\gamma_2\gamma_3\gamma_k$, from which (46.18) at once follows, using (46.9).

X.4. From the Dirac equation:

$$\gamma_k \frac{\partial \psi}{\partial x_k} + \gamma_4 \frac{\partial \psi}{i\,\partial x_0} + M\psi = 0.$$

Taking complex conjugates, transposing and postmultiplying by γ_4:

$$\frac{\partial \psi^\star}{\partial x_k}\gamma_k\gamma_4 - \frac{\partial \psi^\star}{i\,\partial x_0}\gamma_4{}^2 + M\psi^\star\gamma_4 = 0$$

and since $\gamma_k\gamma_4 = -\gamma_4\gamma_k$ and $\bar{\psi} = \psi^\star\gamma_4$, equation (46.21) follows.

X.5. Differentiating (46.24b) and using the two Dirac equations (46.7) and (46.21), (46.25) follows at once.

X.6.

$$(u_r, u_s) = \sum_{a=1}^{4} u_r{}^{(a)*} u_s{}^{(a)}; \;\; \bar{u}_r u_s = \sum_{a,\,b=1}^{4} u_r{}^{(a)*} \gamma_4{}^{ab} u_s{}^{(b)}.$$

The desired results follow directly, using (46.14) and (48.4).

X.7. (*i*) For half the cases this follows from the orthogonality of the spinor parts (*cf* Exercise X.6.); for the remaining cases it follows since the spatial parts of ψ_1 and ψ_2 are orthogonal to those of ψ_3 and ψ_4.

(*ii*) $(\psi_r, \psi_r) = (u_r, u_r) = E_p/mc^2$, from equations (48.4), *i.e.* ψ_r is normalized to a density of E_p/mc^2 particles per unit volume. In the rest system of the particles this reduces to the usual density of 1 particle per unit volume.

(*iii*) The reason for the normalization to a density $\varrho = E_p/mc^2$ particles per unit volume is that this is Lorentz invariant. To see this, consider a volume V containing $\mathcal{N}$ particles having energy E_p. Under a Lorentz transformation to a new coordinate system having a relative velocity $v = \beta c$, $V \to V' = V(1-\beta^2)^{\frac{1}{2}}$, $E_p \to E_p' = E_p(1-\beta^2)^{-\frac{1}{2}}$. Hence

$$\varrho' = \frac{\mathcal{N}}{V'} = \frac{\mathcal{N}}{V}\frac{V}{V'} = \frac{\mathcal{N}}{V}\frac{E_p'}{E_p} = \frac{E_p}{mc^2}\frac{E_p'}{E_p} = \frac{E_p'}{mc^2}.$$

X.8. We must prove $[\hat{\boldsymbol{\sigma}}\mathbf{p}, H] = 0$. This follows at once from equation (47.8), which gives

$$[\hat{\boldsymbol{\sigma}}\mathbf{p}, H] = 2c\gamma_4\mathbf{p}(\boldsymbol{\gamma} \wedge \mathbf{p}) = 0.$$

X.9. If $x_\lambda \to x_\mu' = a_{\mu\lambda}x_\lambda$ and $x_\mu' \to x_\nu'' = b_{\nu\mu}x_\mu'$, then $x_\lambda \to x_\nu'' = c_{\nu\lambda}x_\lambda$, where $c_{\nu\lambda} = b_{\nu\mu}a_{\mu\lambda}$. Then

$$c_{\nu\sigma}c_{\nu\tau} = b_{\nu\mu}a_{\mu\sigma}b_{\nu\lambda}a_{\lambda\tau} = \delta_{\mu\lambda}a_{\mu\sigma}a_{\lambda\tau} = \delta_{\sigma\tau}.$$

Also $|c_{\nu\lambda}| = |b_{\nu\mu}||a_{\mu\lambda}| = 1$, showing that the composite transformation is also a proper Lorentz transformation. From $c_{\nu\lambda} = b_{\nu\mu}a_{\mu\lambda}$ and (49.10) one sees that c_{kl} and c_{44} are real, c_{k4} and c_{4k} are pure imaginary. For in the former case the factors a and b are both real or both imaginary. In the latter case, one factor is always real but the other imaginary.

APPENDIX A: VECTOR ANALYSIS

IN this appendix we give a brief summary of the definitions and theorems of vector analysis required in this book. The reader not familiar with these will find clear and simple expositions in the books by Abraham-Becker[6] or Hague[7].

(1) If **A** is a vector with components (A_x, A_y, A_z) we define its magnitude

$$A = |\mathbf{A}| = \sqrt{\{A_x^2 + A_y^2 + A_z^2\}}. \tag{A 1}$$

(2) If **A** and **B** are vectors, components (A_x, A_y, A_z) and (B_x, B_y, B_z) respectively, we define

(*a*) the scalar product

$$\mathbf{AB} = A_xB_x + A_yB_y + A_zB_z = AB\cos\theta, \tag{A 2}$$

where θ is the angle between the vectors **A** and **B**; in particular $\mathbf{AB} = 0$, if **A** is orthogonal to **B**;

(*b*) the vector product $\mathbf{A}\wedge\mathbf{B}$ as the vector **C** with components

$$\left.\begin{aligned} C_x &= A_yB_z - A_zB_y \\ C_y &= A_zB_x - A_xB_z \\ C_z &= A_xB_y - A_yB_x. \end{aligned}\right\}$$

If **i**, **j**, **k** are unit vectors in the directions of the coordinate axes, so that $\mathbf{A} = \mathbf{i}A_x + \mathbf{j}A_y + \mathbf{k}A_z$, and

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = 1, \quad \mathbf{ij} = \mathbf{jk} = \mathbf{ki} = 0,$$

we can write the vector product as a determinant

$$\mathbf{C} = \mathbf{A}\wedge\mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}. \tag{A 3}$$

C is a vector of magnitude $AB\sin\theta$, whose direction is normal to the plane containing **A** and **B**, such that **A**, **B** and **C** form a right-handed coordinate system.

(3) Let $\varphi(\mathbf{r})$, $\psi(\mathbf{r})$, $\chi(\mathbf{r})$, ... be scalar fields (*e.g.* temperature, electrostatic potential, *etc*), and let $\mathbf{F}(\mathbf{r})$, $\mathbf{G}(\mathbf{r})$, ... be vector fields (*e.g.* fields of force), we then define

(*a*) the gradient of $\varphi(\mathbf{r})$ as the vector

$$\nabla\varphi = \quad \text{grad}\,\varphi = \mathbf{i}\frac{\partial\varphi}{\partial x} + \mathbf{j}\frac{\partial\varphi}{\partial y} + \mathbf{k}\frac{\partial\varphi}{\partial z}, \tag{A 4}$$

(*b*) the divergence of $\mathbf{F}(\mathbf{r})$ as the scalar

$$\nabla F = \quad \text{div}\,\mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}, \tag{A 5}$$

(*c*) the curl of $\mathbf{F}(\mathbf{r})$ as the vector

$$\operatorname{curl} \mathbf{F}(\mathbf{r}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix}.$$

If, in particular, $\mathbf{F} = \operatorname{grad} \varphi$, we obtain

$$\operatorname{div} \operatorname{grad} \varphi \equiv \nabla^2 \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2}; \qquad \text{(A 6)}$$

if $\mathbf{F} = \varphi \mathbf{G}$, then

$$\operatorname{div} (\varphi \mathbf{G}) = \varphi \operatorname{div} \mathbf{G} + \mathbf{G} \operatorname{grad} \varphi; \qquad \text{(A 7)}$$

if $\mathbf{G} = \operatorname{grad} \psi$

$$\operatorname{div} (\varphi \operatorname{grad} \psi) = \varphi \nabla^2 \psi + \operatorname{grad} \varphi \operatorname{grad} \psi. \qquad \text{(A 8)}$$

(4) We state Gauss's theorem for a vector field $\mathbf{F}(\mathbf{r})$:

$$\int_V (\operatorname{div} \mathbf{F})\, d\tau = \int_S F_n dS. \qquad \text{(A 9)}$$

Here S is the surface bounding the volume V. The left-hand integral is a volume integral over V, the right-hand integral is a surface integral over S, dS being a surface element of S and F_n the component of the vector $\mathbf{F}$ in the direction of the outward normal to dS.

If we take $\mathbf{F} = \varphi \operatorname{grad} \psi$, then $F_n = \varphi \partial \psi / \partial n$, where $\partial / \partial n$ denotes differentiation with respect to the outward normal, and Gauss's theorem assumes the form

$$\int_S \varphi \frac{\partial \psi}{\partial n} dS = \int_V \varphi \nabla^2 \psi d\tau + \int_V \operatorname{grad} \varphi \operatorname{grad} \psi d\tau. \qquad \text{(A 10)}$$

Interchanging φ and ψ in this equation and subtracting the resulting equation from (A 10), we obtain

$$\int_S \left(\varphi \frac{\partial \psi}{\partial n} - \psi \frac{\partial \varphi}{\partial n} \right) dS = \int_V (\varphi \nabla^2 \psi - \psi \nabla^2 \varphi)\, d\tau. \qquad \text{(A 11)}$$

Equations (A 10) and (A 11) are known as Green's theorem.

BIBLIOGRAPHY

THE following bibliography, which is clearly in no way exhaustive, is merely a list of books helpful to the understanding of this book or suggestions for further reading.

A. Mathematics (eigenvalue problems, special functions, etc)

1 CHURCHILL, R. V. *Fourier Series and Boundary Value Problems* McGraw-Hill, 1941. (Eigenvalue problems, Fourier series and integrals, Legendre and Bessel functions.)

2 COURANT, R. *Differential and Integral Calculus* Blackie, 1942. (Vol. 1, chap. IX: Fourier series; vol. 2, chap. IV: Fourier integrals.)

Mathematically advanced comprehensive treatments are the following:

3 COURANT, R. and HILBERT, D. *Methoden der mathematischen Physik* vol. 1, Springer, 1931 (and Interscience, 1943). There exists now an English translation, with additions, of this work: *Methods of Mathematical Physics* vol. 1, Interscience, 1953.

4 SOMMERFELD, A. *Partial Differential Equations in Physics* Academic Press, 1949.

5 VON NEUMANN, J. *Mathematische Grundlagen der Quantenmechanik* Springer, 1932 (and Dover, 1943). English translation by R. T. BEYER, Princeton University Press, 1955. (A rigorous abstract treatment of the mathematical scheme of quantum mechanics and its interpretation.)

B. Vector analysis

6 ABRAHAM, M. and BECKER, R. *Classical Theory of Electricity and Magnetism* Blackie, 1946 (chapters 1 and 2).

7 HAGUE, B. *An Introduction to Vector Analysis* Methuen, 1950.

C. Quantum mechanics

8 BORN, M. *Atomic Physics* 3rd ed., Blackie, 1944.

9 HEITLER, W. *Elementary Wave Mechanics* Oxford University Press, 1945. (An elementary introduction to the physical interpretation of wave mechanics.)

10 MOTT, N. F. and SNEDDON, I. N. *Wave Mechanics and its Applications* Oxford University Press, 1948.

11 PAULING, L. and WILSON, E. B. *Introduction to Quantum Mechanics* McGraw-Hill, 1935.

12 SCHIFF, L. I. *Quantum Mechanics* McGraw-Hill, 1949. (This book, which contains many applications, approaches quantum mechanics from a viewpoint rather similar to our own.)

13 SOMMERFELD, A. *Atombau und Spektrallinien* vol. 2, 2nd ed., F. Vieweg, Brunswick, 1944.

[14] Sommerfeld, A. *Wave Mechanics* Methuen, 1930. (This is the English translation of the first edition (1929) of Ref. 13. Hence 13 is much more up-to-date and gives a better and more detailed treatment of many topics.)

The following are more advanced treatises on quantum mechanics:

[15] Dirac, P. A. M. *The Principles of Quantum Mechanics* 3rd ed., Oxford University Press, 1947.

[16] Heisenberg, W. *Physical Principles of the Quantum Theory* University of Chicago Press, 1930.

[17] Pauli, W. *Die allgemeinen Prinzipien der Quantenmechanik, Handbuch der Physik* vol. 24, part 1, Springer, 1933.

[18] Von Neumann, J. See reference 5.

D. Applications of quantum mechanics to particular fields

[19] Bethe, H. A. and Morrison, P. *Elementary Nuclear Theory*, 2nd ed., Wiley, 1956.

[20] Blatt, J. M., and Weisskopf, V. F. *Theoretical Nuclear Physics* Wiley, 1952.

[21] Condon, E. U. and Shortley, G. H. *Theory of Atomic Spectra* Cambridge University Press, 1951.

[22] Mott, N. F. and Massey, H. S. W. *Theory of Atomic Collisions* Oxford University Press, 1949.

[23] Van der Waerden, B. L. *Die gruppentheoretische Methode in der Quantenmechanik* Springer, 1932 (and Edwards, 1944).

[24] Wentzel, G. *Wellenmechanik der Stoss- und Strahlungsvorgänge, Handbuch der Physik* vol. 24, part 1, Springer, 1933 (and Edwards, 1943).

[25] Massey, H. S. W. *Theory of Atomic Collisions, Handbuch der Physik* vol. xxxvi (Atoms II), Springer, 1956.

INDEX*

Adjoint, 72–73, 95
Allowed transitions, 201
Angular momentum, **14**, **25**, *see also* Spin
  addition of, **27**, 128–130, 245–247
  classically, 54
  conservation of, 81, 182
  operators, 34, 54, 104, 105, 229
    commutation relations for, 79, 84, 89, 104, 105, 130, 237–238, 239, 244–245
    eigenfunctions and eigenvalues of, 54–55, 67–68, 104, 105–108, 130
  total, 116, 121
Associated Legendre functions, 50–53

Basis, *see* Matrix representation, Group representation
Bessel functions, 166–168
Bohr magneton, 141
Born approximation, **34**, 160, 175, 252
Boson, 127
Bra vector, 102

Centre-of-mass coordinate system, **35**
Clebsch–Gordan coefficient, 120
Commutation relations, 79, 84, 86, 89, 104
Commutator, 72, 77
Commuting observables (or operators), 77–78
  complete set of, 80, 81, 115–116
Compatibility of several observables, 75–78, 82
Complete orthonormal set, 4, 10
Compton wave-length, 204
Constant of the motion, 69, 81, 101, 115–116, 126, 182, 185, 186
Coulomb scattering, 151–153
Cross-section, 144, 145, *see also* Scattering, Born approximation
Current-density vector, 31, 35, 209

Delta function, 17–19, 145
Density of states, 42–43, 148
Dirac equation
  for free particle, **46**
  for particle in electro-magnetic field, 210–212
  negative energy states, 212, 216
  non-relativistic limit, 211–212, 213–215
  plane-wave solutions, **48**
  relativistic invariance, **49**

Eigenfunctions, 6, 12, 27, 32, 62–63
  complete set of simultaneous, 75–78, 100
  completeness of, 12–13, 34, 62–63
  degenerate, 13–14, 63
  expansion in, 13, 17, 34, 62–63
  interpretation of, 27, 55, 66–68
  linearly independent, 13
  orthogonality of, 9, 33–34, 40, 63, 100, 102, 238–239
Eigenvalue, 6, 12, 27, 32, 62–63
  degenerate, 13–14, 100
  interpretation of, 27, 55, 63, 66–68
  problem, 6, 12, 27, 32
    in matrix form, 97–100, 103, 239–240
  reality of, 33, 63, 100, 102, 238
  spectrum
    continuous, **4**, **17**, 62
    discrete, 14, 32–33, 62
    mixed, 16
Electro-magnetic transitions, 197–200
Electron
  magnetic moment, **26**, **47**
  spin, **26**, **47**, 204
Exchange integral, 139
  operators, 126, 184
Exclusion principle, 127
Expectation value, 63, 66–68, 71, 81, 115

Fermion, 127
Forbidden transitions, 201
Four-vector, 220
Fourier integral, 15, 19, 20, 87
  series, 7, 20
Function
  even and odd, 8, 39
  length of, 9
  magnitude of, 9
  space, 10, 11, 60, 76–77
    complex, 11
    dimensionality of, 10
    manifold of, 13
    product, 113–115, 116–117, 183
    subspace of, 13

Gauss's theorem, 262
Green's theorem, 262
Ground state of
  alpha particle, 129–130
  deuteron, 201, 256
  ^{3}He, 130
  helium atom, 57, 136, 234
  hydrogen atom, 57, 233
Group, 178
  Abelian, 179, 189, 201, 255–256
  axial rotation, 178–179, 189–190
  inversion (or reflection), 179, 188
  Lorentz, 218
  permutation, 179, 189, 201, 255
  representation, **43**, 187
    basis for, 187
    decomposition of, 187
    equivalent, 187
    irreducible, 187, 202
    product, 189
    reducible, 187–188
    reduction of, 187

* The numbers in heavy type refer to the sections

Group-*contd.*
  rotation, **40**, 179, 189
  transformations (or operators), of, **39**, 178
    invariants of, **41**, 186
    unitary, 180, 201, 256
  velocity, 30, 35, 230

Hamiltonian function, 27, 59, 64
  invariance under
    inversion, 56, 182, 185, 186, 193
    permutations, 125, 137, 177, 194
    rotations, 105, 125, 177, 181, 182, 185, 186, 194
  spherically symmetric, 47, 54, 68, 104, 105, 115–116, 125, 140, 141, 181, 182
Hamilton's equations, 59, 88, 234
Hankel functions, 166-168
Heisenberg, *see also* Uncertainty principle
  equations of motion, 101, 103, 240
  picture or representation, 100–101
Hermite polynomials, 45–46
Hermitian, 62–63, 72–73, 95
Hilbert space, 11, 60, 76–77

Identical particles, **28**, 130, 137–140, 160–161, 247–248
Inner product, 2, 4, 9, 11, 91, 114
  invariance under unitary transformations, 73
Intercomhination lines, 198
Invariance, *see* Hamiltonian function
Invariants, *see* Group of transformations
Inversion, 39, 56, 177, 182, 185, 217, 231

Ket vector, 102
Klein–Gordon equation, 203
Kronecker δ symbol, 3

Laboratory coordinate system, **35**
Laporte's rule, 198
Legendre polynomials, 50–51
Life–time, 86
Linear independence, 4, 11
Lorentz transformation, **49**

Magnetic moment, **47**, 111–112
Manifold, 4, 13, 76–77, 79–80, 98, 100
  decomposition of, 187
  invariant, 186
  irreducible, 187
  reducible, 187, 188
  reduction of, 187
Matrix, 92
  diagonal, 94, 102
  eigenvectors and eigenvalues of, 97–100, 102
  multiplication, 91, 93–94
  representation, **21**, *see also* Group representation
    basis of, 90
    change of basis of, 95–97
    eigenvalue problem in, 97–100
    of an operator, 92
    of a vector, 91
  transposed, 95
  unit, 94
Mean square deviation, 67–68, 82
Measurement, 65–70
  complete (maximal), 80
  one observable, of, 65–68, 70–71
  several commuting observables, of, **19**
  several non-commuting observables, of, **20**
  state of system, of, 79–80
  state of system after, 68–70

Non-commuting operators, *see* Operators
Norm, 1, 4, 5, 9, 12
Normalized, 2, 9
Nucleon, 129

Observable, 25–28, 61–63
Operator, 26–28, 61-63
  commuting with the Hamiltonian, 80–81, 101, 115–116, 126
  equations of motion of, *see* Heisenberg
  inverse, 74
  unit, 73
Operators
  commutability of, 77-78, 100
  complete set of commuting, 80
  functions of, **18**, 78
  non-commuting, 82
Optical theorem, 172
Orthogonal, 3, 4, 9
Orthonormal, 3, 9

Parity, 39–40, 45, 51, 53, 56, 88, 182, 185, 193, 230–231
Partial wave analysis, **38**
Pauli
  matrices, *see* Spin matrices
  principle, *see* Exclusion principle
Permutation operator, *see* Exchange operators
Perturbation theory
  time-dependent, **37**
  time-independent, Chapter VII (131–141)
    applications, **31**
    first-order degenerate, **30**
    first-order non-degenerate, **29**
    second-order non-degenerate, 141, 249–250
Phase factor, 28, 60, 61, 108, 109, 120, 133, 243–244, 250
  velocity, 29
Phase-shift, 171
Planck's constant, 22, 23
Plane wave, **10**, 23, 28, 29
Positron, 216
Probability current, 31, 35
  interpretation of quantum mechanics, 28–29, 65–68, 70–71, 115
Pseudo-scalar, 221
Pseudo-vector, 182-183, 221, 222

Quantum number
  angular momentum, 55
  good, 182, 198
  magnetic, 55, 141

Radiative transitions, 197–200
Reduced mass, 143, 157, 232
Relativistic invariance, **49**